JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

3/2001

Special issue edited by Tadeusz Antoni Wysocki

50 YEARS
National Institute of Telecommunications

Editorial Board

Editorial Advisory Board

ISSN 1509-4553

JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

Preface

In the first year of the twenty first century, we are witnessing unprecedented growth in both quality and quantity of services offered by communication systems. While communications practitioners are endeavouring to translate existing recent advancements into product devel-opment, researchers continue scanning the horizon for new ideas and solutions in order to approach theoretical performance predictions. Most of those recent advancements in com-munication systems performance have been only possible because of signal processing ap-plied in all areas of communication systems development and implementation. The special issue *Signal Processing for Communication Systems* consists of an invited paper and eight regular contributions.

In the invited paper *An introduction to envelope constrained filter design*, Cantoni, Vo, and Teo provide an overview of key ideas related to robust continuous time envelope constrained filter design. The paper is concerned with time domain synthesis of a filter whose response to a specified input signal stays within prescribed bounds and in addition has minimal noise enhancement. The specification of the output constraint envelope can arise either from standards set by certain regulatory bodies or practical design considerations.

The second paper *Source enhanced linear prediction of speech incorporating simultaneously masked spectral weighting* by Lukasiak and Burnett deals with a modification of the method to calculate linear predictor coefficients based on the simultaneous masking property of a human ear. It is shown that by applying the proposed modification an improvement in subjective quality can be achieved without an increase of transmission rate at the price of a modest increase in computational complexity.

The next four papers consider different aspects of signal processing that can be applied to emerging wireless and access systems. In the first of these papers Walsh-chirp sequences for wireless applications Wysocki, Wysocki, and Zepernick, describe the new class of poly-phase spreading sequences having a potential for reducing multi-access interference in direct sequence code division multiple access (CDMA) wireless data networks. The second of these papers Memory truncation and crosstalk cancellation for efficient Viterbi detection in FDMA systems by Mertins introduces the new method to design transmission systems for frequency division multiple access (FDMA) over frequency selective channels. The proposed method can be successfully applied to discrete multitone modulation (DMT) or to orthogonal frequency division multiplexing (OFDM) as well as to CDMA. Then, Bourdel, Campo, Melet, and Andrieux present a design path for a chipset to be used in wireless data CDMA transceiver in their paper *From modelling of a CDMA transceiver in indoor environment to an ASIC circuit synthesis*. The series concludes with the paper *Simulating capture behaviour in 802.11 radio modems* where Ware, Wysocki, and Chicharo introduce a new model describing the operation of an 802.11 receiver. That new model is able to reflect the fairness characteristics obtained with an IEEE 802.11 radio modem more accurately than the previous capture models.

JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

The next two papers *Error statistics for concatenated systems on non-renewal time-varying channels* by Pimentel and *Stop criteria for retransmission termination in soft-combining algorithms* by Zepernick and Caldera are concerned with improving transmission quality by applying different error-control strategies.

The final paper of the issue *Precise measurement of complex permittivity of materials for telecommunications devices* by Nakamura and Nikawa is concerned with application of signal processing for measuring parameters of materials used for design of telecommunication devices.

The Guest Editor wish to thank the numerous reviewers who have volunteered their precious time in order to provide feedback to the authors and to improve quality of the issue. He would also like to express his gratitude to the contributors and to the Editorial Office of the Journal for the valued assistance in bringing this issue to fruition.

Tadeusz Antoni Wysocki
Guest Editor

Invited paper

An introduction to envelope constrained filter design

Antonio Cantoni, Ba-Ngu Vo, and Kok Lay Teo

Abstract — Envelope constrained filter design is concerned with the time domain synthesis of a filter whose response to a specified input signal stays within prescribed upper and lower bounds and in addition has minimal noise enhancement. In many practical applications, a "soft" approach, such as least mean square, is not the most suitable and it becomes necessary to use "hard" constraints such as the ones considered in the paper. We present an overview of key ideas related to robust continuous time envelope constrained filter design.

Keywords — filter, optimisation, constraints, envelopes, time domain, noise, robustness.

1. Introduction

In the continuous time envelope constrained (EC) filter design problem, we consider the design of a filter such that the noiseless response to a specified excitation fits into a prescribed envelope. Furthermore, we seek those filters that minimize a cost functional that is appropriate for the application of the filter.

Often in a time domain filter synthesis problem, the performance criterion used is the mean square error between the filter output and some desired signal. However, in many practical applications, this "soft" approach is not the most suitable and it becomes necessary to use "hard" constraints such as envelope constraints. Moreover, problems often arise in practice where it is crucial that the shaped signal fits into a prescribed envelope. The specifications of the output constraint envelope can arise either from standards set by certain regulatory bodies or practical design considerations. In telecommunications systems, pulse shapes used in transmission systems are specified by recommendations issued by standards bodies [1 – 4]. In sonar, radar ranging systems and imaging systems, the constraints arise from consideration of resolution and are associated with pulse compression technique requirements [5 – 8].

Over the past twenty years, the work on envelope constrained filters has evolved to take into account many engineering issues. In our overview, we present the key ideas related to robust continuous time envelope constrained filter design. Numerical methods for obtaining the solution to the problems formulated are not considered in this paper. The reader is referred to the following references for a range of efficient techniques for computing the optimal solution [9 – 18].

2. The envelope constrained filter applications

In this section, we briefly review a few applications that have motivated the study of envelope constrained filter design. The applications covered are by no means exhaustive and are intended to bring out a number of different aspects of the envelope constrained filter design problem.

2.1. Radar application

In radar and sonar, narrow pulses are required for range resolution and clutter reduction, but it is also important that the transmitted pulses contain sufficient energy for long-range detection. These two conflicting requirements would seem to dictate narrow pulses with high peak power. However, because of transmitter design considerations, peak power is limited, pulses of relatively long duration are transmitted and an operation known as **pulse compression** is performed at the receiver as illustrated in Fig. 1. This compression is most commonly achieved with a **matched filter**, that is, by correlating the incoming signal with the time-shifted and/or frequency shifted copies of the transmitted waveform. Matched filtering is well known to be optimal with respect to various performance criteria.

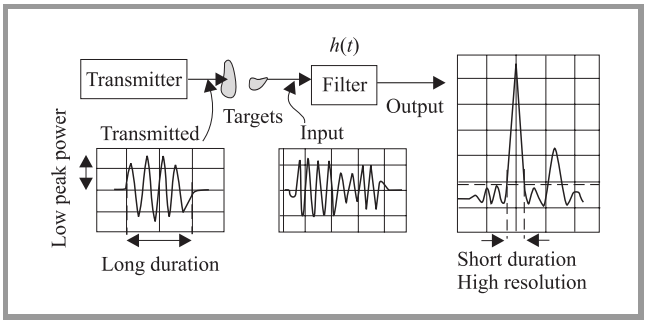


Fig. 1. Pulse compression.

The presence of a signal is detected when the matched filter output exceeds a threshold value, and parameters such as time delay and (Doppler) frequency shift are **estimated** by locating the peak output in time and frequency. If a signal s, non-zero on the interval $(t_0, t_0 + T)$, appears at the input to a pulse-compression filter, the output typically consists of a main peak surrounded by sidelobes, as shown in Fig. 1. When a radar must distinguish among multiple targets, these sidelobes can cause false detections and impede resolution of adjacent pulses. For example, if the transmitted pulse is a 13-bit Barker code and a matched filter

receiver is used, the output is compressed to 1/13 of the original pulse length, improving range resolution with no detection loss. Unfortunately, the output pulse is accompanied by 12 sidelobes, where the amplitude of each of these sidelobes is 1/13 of that of the main lobe peak. Thus, it is necessary in certain cases to remove or at least restrict the height of these sidelobes to an acceptable level. Of course there will be some noise penalty that results.

Problems of this type are often treated by using a least square approach, where the mean squared difference between the output and some desired pulse shape is minimized. If the sidelobes are reduced using a least square cost, there is no guarantee that the appearance of low-energy high-peak sidelobes (which can cause false detections) can be avoided. Moreover, the solution can be sensitive to the detailed structure of the desired pulse, and it is usually not obvious how the shape of the desired pulse should be altered in order to achieve better performance in terms of low level sidelobes.

The reduction of sidelobes can be formulated in terms of an envelope constrained problem: find the filter (if it exists) that causes least detection loss while reducing the output sidelobes to a specified fraction of the mainlobe peak. The envelope or pulse shape mask takes on the form shown in Fig. 2. Since the probability of detection is proportional to the signal to noise ratio, which in turn, is inversely proportional to the square of the norm of the filter, we are faced with the task of finding a minimum norm filter subject to the constraint that the sidelobe peaks remain within the specified values.

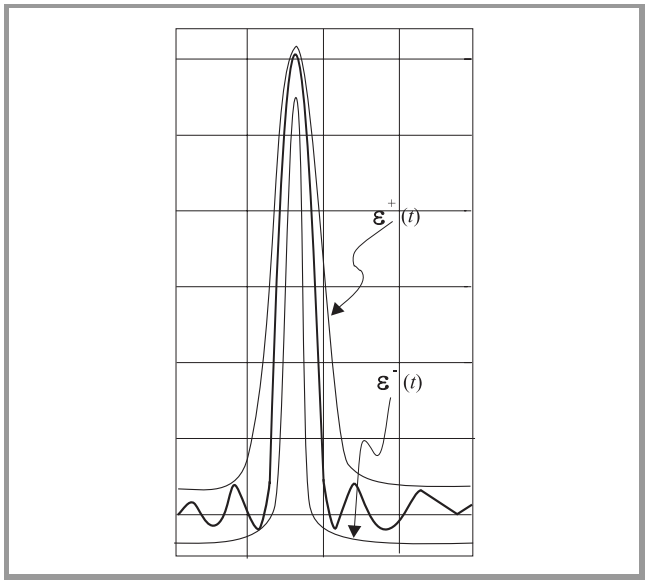


Fig. 2. Pulse shape constraints for radar/sonar problem.

2.2. Digital transmission application

In telecommunication standards, the performance of a digital link is often characterized by pulse masks applied to the test pulses (see, for example, the CCITT recommendations [1, 2]). The signal s would correspond to the test signal specified in the standards. As an example, consider the equalization of an RG59B/U coaxial cable channel whose attenuation follows an approximate $\sqrt{f}$ law (one with a 150 m length has approximately 12 dB loss at 70 MHz). The frequency response of a coaxial cable of length l is given by

$$H(j\omega) = e^{-Al\sqrt{j\omega}} = e^{-Al\left(\sqrt{\omega/2}+j\sqrt{\omega/2}\right)} .$$

Let A_0 denote the attenuation of the cable in dB at a frequency f_0 Hz. Then,

$$A_0 = (20Al\log e)\sqrt{\frac{\omega_0}{2}} = (20Al\log e)\sqrt{\pi f_0} ,$$

where $\omega_0 = 2\pi f_0$. The impulse response of the cable is given by

$$h(t) = \frac{Al}{2\sqrt{\pi t^3}} \exp\left(-\frac{A^2l^2}{4t}\right), \; t > 0 .$$

A plot showing the impulse response for several values of Al is given in Fig. 3. Since Al increases with increasing cable length, the peak of the pulse in transit decreases and its base width widens as the cable lengthens. In order to successfully detect these pulses, the width must be compressed by means of pulse shaping networks at the receiving end.

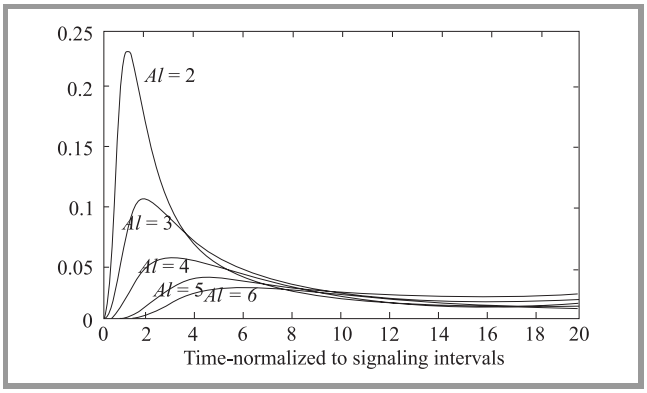


Fig. 3. Impulse response of coaxial cable for various lengths.

For a digital transmission channel consisting of a coaxial cable operating at the DX3 rate, i.e. 45 Mb/s (see [1, 2]), the American National Standards Institute (ANSI) specifies that at the pulse received, after transmission should fit in the mask illustrated in Fig. 4. From the characteristics of the cable shown in Fig. 3 it can be seen that for long cables the received pulse will not fit in the mask. An equalizing filter is required to shape the impulse response of the cable so that it fits in the envelope given by the DSX3 pulse template in Fig. 4. In Fig. 4 we have also shown the received pulse and the pulse at the output of an optimal EC filter. The cable has a 30 dB attenuation at frequency $2\pi/\beta$, where β denotes the baud interval ($22.35 \cdot 10^{-9}$ s).

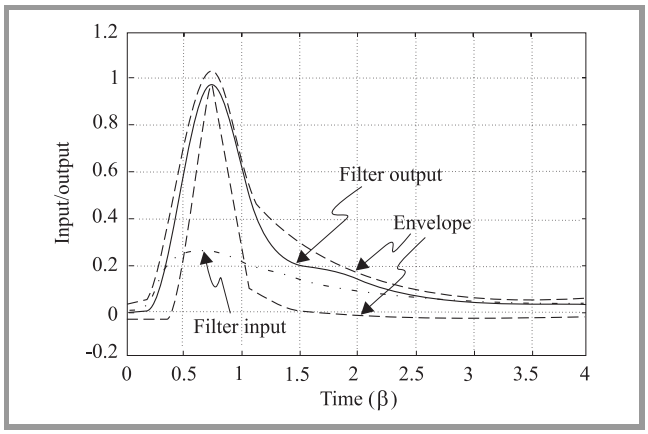


Fig. 4. DSX3 pulse template, coaxial cable and filter output.

Consider another transmission system example shown in Fig. 5, where a pulse pre-shaping filter is placed before the channel. When a rectangular pulse is transmitted, the noiseless output of the channel is required to fit into a T1 mask specified by the ANSI. The T1 baud rate is 1.544 Mb/s. It makes sense to minimize the norm of the pulse shaping filter to reduce the crosstalk to other signal as this is proportional to the transmit signal power.

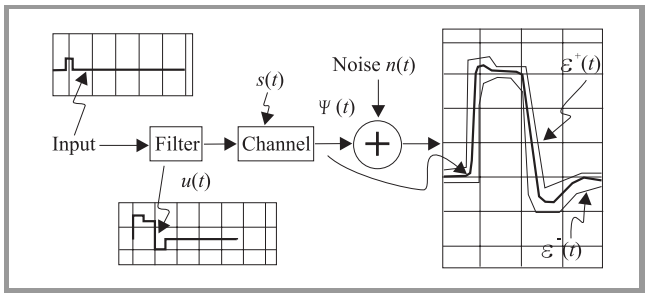


Fig. 5. Pre-shaping of pulse.

3. Formulation of optimal envelope constrained filters

In this section, we first introduce the basic EC filter design problem. This includes consideration of possible cost functionals for selecting the optimum filter. Then we expand the basic formulation by considering the problem of designing EC filters that are robust to signal modeling errors and filter implementation errors.

3.1. Basic envelope constrained filter design problem

We now formalize somewhat the definition of the EC filtering problem in its simplest form. The previous section has outlined some applications that motivated the study of the EC filtering problem without an explicit problem statement. This section presents a more precise statement that covers all the applications previously discussed for both analog and hybrid filters.

3.1.1. Analog filters

Consider the filtering function shown in Fig. 6 be it for pulse compression or equalization. The excitation s entering the filter is corrupted by additive zero-mean, stationary noise n. The impulse response u of the linear time-invariant

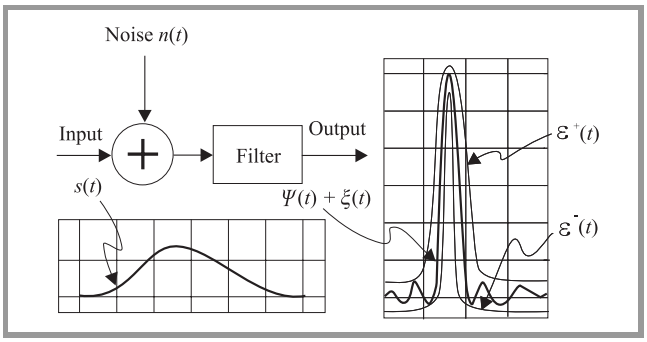


Fig. 6. Receiver model and output mask.

filter is to be determined and is restricted to be Lebesque square-integrable on $[0,\infty)$. The output consists of two components ψ and ξ due to the signal and noise respectively. The noiseless output ψ is given by the convolution $s * u$:

$$\psi(t) = (s*u)(t) = \int_0^\infty s(t-\lambda)\,u(\lambda)\,d\lambda\,.$$

The notation Ξ_s is adopted to denote the mapping of u to the filter response $\psi = s*u$, i.e.

$$\Xi_s u = s*u\,. \tag{1}$$

Let ε^- and ε^+ be two piece-wise continuous functions of time representing the lower and upper boundaries of the output mask respectively. Then the envelope constraints require the filter output ψ to fit into a fully specified envelope as depicted in Fig. 6, i.e.

$$\varepsilon^-(t) \le \psi(t) \le \varepsilon^+(t), \quad \forall t \in \Omega \subset [0,\infty)\,. \tag{2}$$

Define

$$d = 0.5(\varepsilon^+ + \varepsilon^-) \text{ and } \varepsilon = 0.5(\varepsilon^+ - \varepsilon^-)\,.$$

Then (2) can be written as

$$|\Xi_s u - d| \le \varepsilon\,.$$

So far we have defined the feasible region for the filters we wish to find. In many cases, the feasible region is not empty and there are many filters for which (2) holds. However, in many applications, there are other considerations that in fact lead us to select a specific filter. For example, it may be desirable to minimize the noise enhancement of the filter if the signal input to filter is corrupted by noise or it may

be necessary to minimize the filter output power to contain crosstalk. We consider in detail the noisy input case and define our cost function f to be the output noise power. Let ξ denote the noise component of the filter output. Then,

$$\xi(t) = \int_0^\infty n(t-\lambda)\,u(\lambda)\,d\lambda\,. \tag{3}$$

Assuming that the additive noise n at the receiver input is zero-mean, stationary, with autocorrelation $R_{nn}(\tau)$, the output noise power is:

$$f(u) = E\left[\xi^2(t)\right] = \int_0^\infty \int_0^\infty R_{nn}(\kappa-\lambda)\,u(\lambda)\,u(\kappa)\,d\kappa d\lambda\,. \tag{4}$$

By defining a linear operator L based on the autocorrelation function R_{nn} as

$$(Lu)(\lambda) = \int_0^\infty R_{nn}(\kappa-\lambda)\,u(\kappa)d\kappa \tag{5}$$

and using the usual inner product, the cost functional defined by Eqs. (1) and (4) can be expressed in a more convenient and intuitive form as a quadratic function of the filter impulse response:

$$f(u) = \langle u, Lu\rangle\,. \tag{6}$$

The cost functional is strictly convex when L is positive definite. In this case,

$$\|\,u\,\|_L \equiv \sqrt{\langle u, Lu\rangle}$$

is a norm.
Suppose that in addition to being additive, zero-mean and stationary, the input noise n is also **white**, with autocorrelation

$$R_{nn}(\tau) = N_0\delta(\tau)\,,$$

where δ is a unit impulse. Then, the output noise power $E\,[\xi^2(t)]$, Eq. (4), is proportional to the square of the L_2-norm of the filter,

$$\begin{aligned} E\left[\xi^2(t)\right] &= N_0\int_0^\infty\int_0^\infty \delta(\kappa-\lambda)\,u(\lambda)\,u(\kappa)\,d\kappa d\lambda = \\ &= N_0\int_0^\infty u^2(\lambda)\,d\lambda = N_0\|u\|_2^2\,. \end{aligned} \tag{7}$$

The cost functional in this case is strictly convex (not all norms are strictly convex), and L, defined by Eq. (5), is the identity operator.
In the derivation of the previous cost functionals, it is assumed that the statistics of the noise are known. In some applications, the exact noise statistics may not be known. We now consider one approach for dealing with this case.

Suppose that the noise spectral density is denoted as $\Phi_N(\omega)$. It can be verified that the output noise power due to the input noise n is given by

$$P_N = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_N(\omega)|U(j\omega)|^2 d\omega\,,$$

where U denotes the Laplace transform of u. Assume that Φ_N the noise power density satisfies

$$\|\Phi_N\| = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_N(\omega)\,d\,\omega \le 1$$

but is otherwise unknown. It makes sense to consider the output noise power for the worst case input noise, i.e.

$$P_N = \max_{\|\Phi_N\|\le 1}\frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_N(\omega)|U(j\,\omega)|^2 d\,\omega\,.$$

It can be shown [26], that

$$\|U\|_\infty^2 = \max_{\|\Phi_N\|\le 1}\frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_N(\omega)|U(j\,\omega)|^2 d\omega\,, \tag{8}$$

where $\|U\|_\infty$ denotes the H_∞ -norm of U and is defined as $\|U\|_\infty = \sup_{\omega\in\mathbf{R}}|U(j\omega)|$.
Hence, minimizing the output noise power for the worst case input noise is equivalent to finding a filter of transfer function $U(s)$ with minimum H_∞ -norm. In this case, the cost functional can be expressed explicitly in terms of u as

$$f(u) = \left\| \int_0^\infty u(t)e^{-st}dt \right\|_\infty^2. \tag{9}$$

This cost functional is convex (but not strictly convex). In cases in which the input signal is subject to random disturbance with unknown but bounded power spectrum, the H_∞ optimization approach may offer a robust design.

The use of these and similar convex cost functionals can also be motivated by consideration of the filter's sensitivity to implementation errors and uncertainties in filter parameters.

3.1.2. Hybrid filters

Advances in the development of digital processors motivate the consideration of filter structures realized with digital components. A continuous-time filter can be implemented as a hybrid filter composed of an A/D converter, a discrete-time filter, a D/A converter and a post-filter. The EC design filter problem is to determine the discrete-time component of such a filter so as to minimize the effect of input noise whilst satisfying the constraint that the response of the filter to a specified signal fits into a prescribed mask.
The hybrid filter is shown in Fig. 7. The post-filter's function is to smooth the output of the D/A block. In other words, the combined function of the D/A block and the post-filter is to interpolate the (discrete-time) output of the digital processor. The absence of a post-filter corresponds to piece-wise constant interpolation of the digital output. Next on the ladder of complexity is linear interpolation. The output of the system varies from a staircase to much smoother waveforms depending on the type of interpolation. In practice, the post-filter is often implemented as a lowpass filter with cut-off frequency being half of the sampling frequency, fo example, Butterworth, Bessel, Chebyshev and elliptic filters.

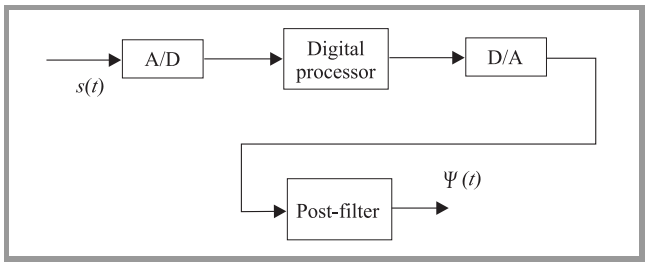


Fig. 7. Hybrid filter.

It is assumed that the incoming signal is sampled at or above the Nyquist rate. To simplify matters, quantization effects inherent in digital processes are neglected. This assumption results in a linear system.

Consider a digital processor that has a discrete-time impulse response $u \in l_2$. Then, the response of the hybrid filter to the continuous-time excitation x is given by

$$\psi(t) = \sum_{i=-\infty}^{\infty} \sum_{j=0}^{\infty} x\big[(i-j)\tau\big] u(j) \Lambda(t - i\tau), t \in [0,\infty) \,, \quad (10)$$

where

$$\Lambda(t) = \int_0^{\infty} \Pi(\lambda) h(t-\lambda) d\lambda \,, \quad (11)$$

$$\Pi(t) = \begin{cases} 1, & t \in [0, \tau] \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

and $h(t)$ is the impulse response of the post-filter.

Assuming appropriate post-filtering such that (10) converges for all $t \in [0,\infty)$ (e.g. bounded input bounded output stability). Then, ψ is bounded and continuous on $[0,\infty)$. Moreover, the mapping defined by

$$(\Xi_x u)(t) = \sum_{i=-\infty}^{\infty} \sum_{j=0}^{\infty} x\big[(i-j)\tau\big] u(j) \Lambda(t - i\tau) \quad (13)$$

is a continuous linear operator. The constraint set for this problem

$$\mathscr{F} = \big\{ u \in l_2 : \varepsilon^- \le \Xi_s u \le \varepsilon^+ \big\}$$

is thus convex. In analog or discrete filtering, the operator Ξ_x corresponds to a convolution by x under appropriate signal spaces.

A simple cost functional that one could use for the EC problem is the norm of the discrete-time filter. However, unlike the discrete-time the output noise power of the hybrid filter is not directly proportional to this norm. Assuming stationary input noise samples $n(i\tau)$, it can be shown that the output noise $\xi(t) \equiv (\Xi_n u)(t)$ is cyclo-stationary with period τ and that an appropriate well-defined cost for the EC problem with hybrid filter is the averaged output noise power given by

$$f(u) \equiv \frac{1}{\tau} \int_0^{\tau} E\big[\xi^2(t)\big] dt = \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} u(l) u(m) L_{l,m} \,, \quad (14)$$

where

$$L_{l,m} = \frac{1}{\tau} \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} R_{nn}\big((j-k)\tau\big) \times$$
$$\times \int_0^{\tau} \Lambda\big[t - (j+l)\tau\big] \Lambda\big[t - (k+m)\tau\big] dt \quad (15)$$

and $R_{nn}\big((j-k)\tau\big)$ is the input noise autocorrelation.

3.2. Envelope constrained filtering problem

The EC filtering problem can be expressed as

$$\begin{aligned} &\min f(u) \\ &\text{subject to} \quad |\Xi_s u - d| \le \varepsilon \,, \end{aligned} \quad (16)$$

where ε^- and ε^+ are two continuous functions of time representing the lower and upper boundaries of the output mask respectively and an appropriate cost functional as discussed above is selected.

To eliminate output envelopes that permit the trivial solution $u = 0$, it is sufficient to assume the existence of an open subset I of Ω such that $\varepsilon^+(t)\varepsilon^-(t) > 0, \forall t \in I$. This means that there is an open interval where both the upper and lower boundaries have the same sign.

3.3. Robust envelope constrained filter design

Assuming that the set of feasible filters does not contain the origin, i.e. no trivial solution, and since we seek feasible filters with smallest possible norm, it follows that the optimum filter always lies on the boundary of the feasible set. This means the response of the optimum filter to the prescribed input touches the output envelope at some points. Consequently, it is to be expected that disturbances in the prescribed input or implementation error can cause the output constraints to be violated. The problem of designing filters that are robust to such disturbances or techniques for providing a guard band on the output mask are essential to the implementation of practical EC filters. Of course, the penalty for robustness is a possible increase in the cost functional.

Three robustness formulations are introduced in the following sections. The first incorporates input uncertainty into the constraints. We refer to this as the EC with uncertain input (ECUI). The second approach incorporates filter implementation uncertainty into the constraints. The third approach does not consider the source of the disturbances but deals with possible disturbances by forcing the filter output away from the envelope boundaries. We refer to this as the constraint robustness problem.

3.4. EC with uncertain input

The EC filtering problem with uncertain input (ECUI), [22], addresses the robustness to input disturbances by allowing for uncertainty in the input pulse (Fig. 8). Here the input s is not specified exactly, but is known to lie within an input envelope described by upper and lower boundaries

s^- and s^+. The filter is required to fit the response of all excitations within the boundaries s^- and s^+ into the output mask. Of course, the penalty for robustness is the increased noise gain of the filter. For example in the radar sidelobe reduction problem, it is desirable to design a sidelobe reduction filter which is robust in the sense that its output sidelobes remains small even if the input signal is slightly different from the nominal input. In channel equalization, the ECUI approach can handle the case where the channel characteristics are not known exactly but are known within limits. Note that the EC filtering problem is a special case of the ECUI problem.

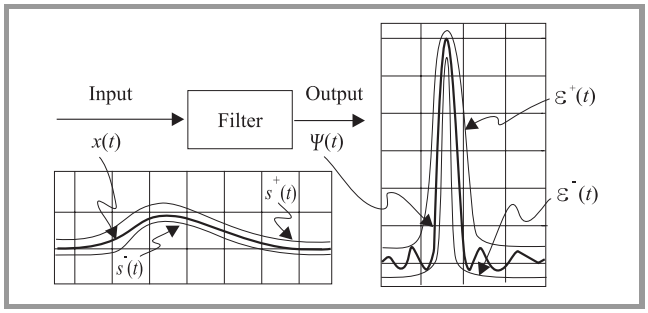


Fig. 8. EC filtering with uncertain input.

When the excitations are also constrained to stay within an envelope described by upper and lower boundaries $s^+, s^- \in L_2[0,\infty)$. The feasible set F is given by

$$\mathscr{F} = \left\{ u \in L_2[0,\infty) : |\Xi_x u - d| \le \varepsilon, \forall x : |x - s| \le \theta \right\} .$$

It can be shown that the feasible region $\mathscr{F}$ is closed and convex. However, while the description of the feasible region in terms of the set of possible inputs is adequate for characterizing the problem it is not useful for computational purposes. To find out if a filter is feasible one would need to compute its response to every signal inside the input mask. There are no standard numerical techniques for handling problems with constraints of this form.
The following result has been established in [27].

Theorem 1: *$|x * u - d| \le \varepsilon, \forall x \in L_2[0,\infty) : |x - s| \le \theta$ if and only if $|s * u - d| + \theta * |u| \le \varepsilon$.*
This result enables us to eliminate the perturbed input x from the original constraint expression and allows problem to be stated as follows:

$$\begin{aligned} &\min f(u) \\ &\text{subject to} \quad G(u) = |\Xi_s u - d| + \Xi_\theta |u| - \varepsilon \le 0 . \end{aligned} \tag{17}$$

The cost functional and the constraint are continuous and convex function of the filter impulse response u. The parameters of this constraint function only involves the known signals d, ε, s and θ. Note that G is not always everywhere differentiable with respect to u.

3.5. Filter implementation uncertainty

In practice, it is often the case that a designed filter or system impulse response cannot be implemented exactly, there are always implementation errors. These errors can arise from component mismatch or quantization of filter coefficients. So we are interested in finding filters which ensure that the envelope constraints are satisfied in the presence of implementation errors (if such filters exist).
It is assumed that for a given filter u, we can achieve actual implementations that are known to be within δ of u, where δ is a bounded function, i.e. any implementation of u belong to the set

$$\mathscr{V}(u,\delta) \equiv \left\{ v \in L_2[0,\infty), |v - u| \le \delta \right\} .$$

In the EC filtering problem we seek filters u whose responses to a prescribed signal s stay within some specified the output envelope. With uncertainty in the implementations, we are interested in those u such that the response of any elements of $\mathscr{V}(u,\delta)$ to the signal s stays inside the output envelope, i.e.

$$|\Xi_s v - d| \le \varepsilon, \forall v : |v - u| \le \delta . \tag{18}$$

Similar to the uncertain input problem, the description of the constraint is not very useful for computational purposes. We can show that the statement in (18) is equivalent to

$$|\Xi_s u - d| + |\Xi_s|\delta \le \varepsilon . \tag{19}$$

Thus, to ensure robustness to implementation errors, we effectively tighten the output mask by reducing ε to $\varepsilon - |\Xi_s|\delta$.

Of course, if the uncertainty δ in the implementation is too large, the resulting output mask would not admit a feasible solution. We can determine from the implementation uncertainty δ whether robustness can be achieved. For instance, if there exists an interval $I \in [0,\infty)$ such that $\left(|\Xi_s|\delta\right)(t) \ge \varepsilon(t)$ for $t \in I$, then there is no feasible solution.

3.6. Constraint robustness

In this section, we present a technique for providing a guard band on the output mask and investigate how the trade-off between the noise gain and the constraint robustness can be achieved by proposing a new optimization problem. This is a generic approach which does not take into consideration the source or cause of the perturbation. Instead, we try to force the filter output to stay as far away from the mask as possible subject to some specified maximum allowable increase on the noise gain (Fig. 9).
For a given filter u (which may or may not satisfy the envelope constraints), consider the difference between its response $\Xi_s u$ and upper or lower mask boundary defined by

$$\left[\phi^+(u)\right](t) = \left(\Xi_s u\right)(t) - \varepsilon^+(t) , \tag{20}$$

$$\left[\phi^-(u)\right](t) = \varepsilon^-(t) - \left(\Xi_s u\right)(t) . \tag{21}$$

It is clear that if $\phi^+(u)$ and $\phi^-(u)$ are non-positive for all $\forall t \in \Omega$ then u satisfies the output constraints. To quantify

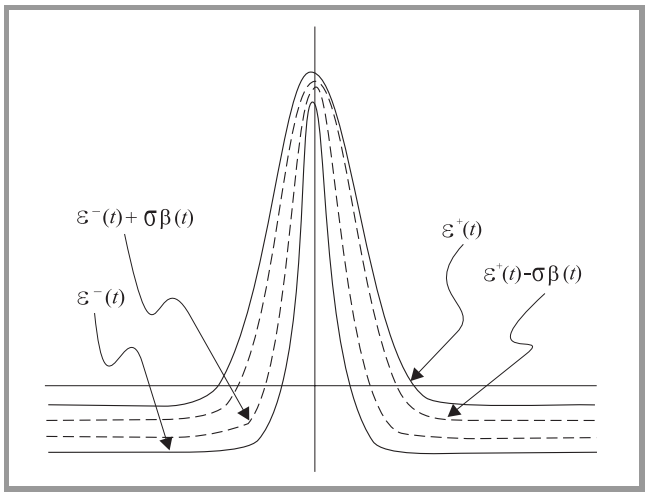


Fig. 9. Mask margins.

the notion of robustness we define its constraint robustness margin as

$$\rho(u) = \min\left\{\min_{t\in\Omega}\left[-\phi^{+}(u)\right](t),\quad \min_{t\in\Omega}\left[-\phi^{-}(u)\right](t)\right\}. \tag{22}$$

The feasible region of the EC filtering problem can now be expressed in terms of the robustness margin as

$$\mathscr{F} = \left\{u \in L_2[0,\infty) : \rho(u) \geq 0\right\}. \tag{23}$$

Moreover, if $\rho(u) > 0$, the minimum distance of the output $(\Xi_s u)(t)$ from the output mask is at least equal to $\rho(u)$. Therefore we may say that the filter u is robust with constraint robustness margin $\rho(u)$.

In practice, it may be necessary to have a larger constraint robustness margin over certain intervals. In this case, β can be used to specify the weightings in different time intervals. Define the weighted constraint robustness as follows:

$$\rho_\beta(u) = \min\left\{\min_{t\in\Omega}\frac{[-\phi^{+}(u)](t)}{\beta(t)},\ \min_{t\in\Omega}\frac{[-\phi^{-}(u)](t)}{\beta(t)}\right\}, \tag{24}$$

where β is a positive, piece-wise continuous function which is normalized so that it attains a minimum of unity on Ω_c and $\delta > 0$ is a constant which specifies the allowable amount of increase of the output noise power in the design of the optimal filter with maximum constraint robustness margin.

Let u^0 denote the optimal solution of the EC filtering problem without any additional robustness constraints, that is the solution to Eq. (16). The EC problem with robustness constraint is the following constrained optimization problem

$$\max \rho_\beta(u)$$
$$\text{subject to } \|u\|^2 \leq (1+\delta)\|u^0\|^2. \tag{25}$$

We can transform this to an equivalent problem as follows:

$$\begin{gathered}\min f(u,\sigma) \equiv -\sigma\\ \text{subject to } g_1(t,u,\sigma) \equiv \sigma\beta(t) - d(t) +\\ -\varepsilon(t) + (\Xi_s u)(t) \leq 0,\\ g_2(t,u,\sigma) \equiv \sigma\beta(t) + d(t) - \varepsilon(t) - (\Xi_s u)(t) \leq 0,\\ g_4(t,u,\sigma) \equiv \|u\|^2 - (1+\delta)\|u^0\|^2 \leq 0.\end{gathered} \tag{26}$$

3.7. Theoretical foundation

From a mathematical point of view the following two theorems can be considered to capture all that one needs to know about the EC filtering problem defined in the previous sections.

Theorem 2: *Let $\mathscr{F}$ be a closed and convex subset in a Hilbert space H, and f a continuous and convex functional on H. If $\mathscr{F}$ is non-empty, then there exists a $u^0 \in \mathscr{F}$ such that*

$$f(u^0) \leq f(u),\ \forall u \in \mathscr{F}.$$

Moreover, if f is strictly convex, then u^0 is unique.

Theorem 3: *Let $\mathscr{F}$ be a closed and convex subset of a separable Hilbert space H with non-empty interior $\mathscr{F}^0$, and f a continuous and convex functional on H. Let u^0 be a minimum of f on $\mathscr{F}$ and u_n^0 a minimum of f on $\mathscr{F} \cap \left[\{v_i\}_{i=0}^n\right]$, where $\left[\{v_i\}_{i=0}^\infty\right]$ denotes the linear span of $\{v_i\}_{i=0}^\infty$. If $\{v_i\}_{i=0}^\infty$ is total in H, then $f(u_n^0)$ monotonically decreases to $f(u^0)$ as n tends to infinity. Moreover, if f is an inner-product-induced norm, then*

$$\lim_{n\to\infty}\|u_n^0 - u^0\| = 0.$$

Theorem 2 tells us about the existence of a solution and its uniqueness. Theorem 3 tells us about the convergence of finite structured filters to the optimal filter.

From a numerical analysts point of view, we would be concerned with the definition and characterization of algorithms for computing a solution. In fact, this aspect has been studied extensively both at a generic level and for the EC problem specifically, [10, 14 – 18, 20]. However, the treatment of these aspects is beyond the scope of this paper.

3.8. Finite filter structures

3.8.1. Analog filters

Suppose that one seeks to realize a filter by designing a network whose impulse response approximates that of the optimum EC filter in some sense, e.g. least squares, Páde, orthogonal approximations. The important question that arises is: would this approximation still satisfy the convex constraints? Thus, it would seem more appropriate to choose a particular filter structure and then impose the constraints. Sub-optimum solutions based on an appropriate choice of basis functions for the filter are likely to be more useful from a practical viewpoint.

Let $\{v_i\}_{i=0}^{\infty}$ be a total sequence in l_2, and let the finite-structured filters have impulse responses given by $K\boldsymbol{a} \in l_2$, where $K : \boldsymbol{R}^n \to l_2$ is a bounded linear operator defined by

$$K\boldsymbol{a} = \sum_{i=0}^{n-1} a_i v_i = \boldsymbol{v}^T \boldsymbol{a} ,$$

where $\boldsymbol{v} = \left[v_0, v_1, \ldots, v_{n-1}\right]^T$.
The feasible region, $F \cap \left[\{v_i\}_{i=0}^{n-1}\right]$, is hence embeded in an n-dimensional sub-space and can be characterized by the set of feasible filter coefficients which is defined in $\boldsymbol{R}^n$ as

$$\left\{\boldsymbol{a} \in \boldsymbol{R}^n : \forall x \in S, x * (K\boldsymbol{a}) \in \Psi\right\} .$$

To obtain an explicit expression for the cost functional in terms of the vector of filter parameters $\boldsymbol{a}$, we have

$$f\big(K(\boldsymbol{a})\big) = \langle K\boldsymbol{a}, LK\boldsymbol{a}\rangle = \boldsymbol{a}^T K^{\dagger} LK\boldsymbol{a} ,$$

where $K^{\dagger} : l_2 \to \boldsymbol{R}^n$ is the adjoint (operator) of K defined by $K^{\dagger}u = \left[\langle v_0, u\rangle, \langle v_1, u\rangle, \ldots, \langle v_{n-1}, u\rangle\right]^T$. The operator $K^{\dagger}LK : \boldsymbol{R}^n \to \boldsymbol{R}^n$ can be represented by the following $n \times n$ matrix, (often called a Gram matrix):

$$K^{\dagger}LK = \begin{bmatrix} \langle v_0, Lv_0\rangle & \langle v_0, Lv_1\rangle & \ldots & \langle v_0, Lv_{n-1}\rangle \\ \langle v_1, Lv_0\rangle & \langle v_1, Lv_1\rangle & \ldots & \langle v_1, Lv_{n-1}\rangle \\ \vdots & & \ddots & \vdots \\ \langle v_{n-1}, Lv_0\rangle & \langle v_{n-1}, Lv_1\rangle & \ldots & \langle v_{n-1}, Lv_{n-1}\rangle \end{bmatrix} . \quad (27)$$

The Gram matrix $K^{\dagger}LK$ is positive semi-definite since L is positive semi-definite. Furthermore if L is positive definite and the v_i's are linearly independent then $K^{\dagger}LK$ is also positive definite. In most practical situations, L is the identity operator and K is specified by a finite subset of a complete set of orthonormal v_i's over the interval Ω_u. In this case it is easily seen that the Gram matrix $K^{\dagger}K$ is the identity matrix I. The main reason for considering filter realization by orthonormal sets is that it is possible to generate and combine certain set of vectors using simple finite filter structures. The most notable examples are Laguerre and Legendre functions.
There are two main techniques for building filters. The parallel structure is the simplest and most direct method (Fig. 10). The transfer function for each block is denoted by $\Theta_i(s), i = 0, 1, \ldots, n-1$ and the transfer function of the filter is immediately given by

$$U(s) = \sum_{i=0}^{n-1} a_i \Theta_i(s) .$$

In this case, $v_i = \theta_i$, where θ_i is the inverse Laplace transform of Θ_i. The parallel structure is widely used for many filters employing orthonormal functions.
The second method is the transversal structure (Fig. 11). The transfer function for each block is denoted by

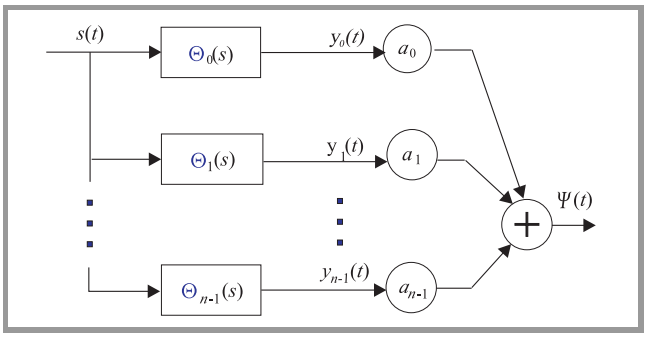


Fig. 10. Parallel filter structure.

$\Theta_i(s), i = 0, 1, \ldots, n-1$. Let V_i denote the Laplace transform of v_i. Then

$$V_i(s) = \prod_{j=0}^{i} \Theta_i(s)$$

and the transfer function of the transversal filter is given by

$$U(s) = \sum_{i=0}^{n-1} a_i \prod_{j=0}^{i} \Theta_i(s) .$$

The best known examples of filters employing the transversal structure are Legendre or Laguerre filters or when the Θ_i's are pure delay elements.

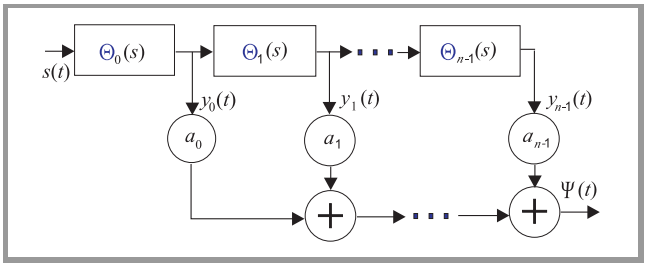


Fig. 11. Transversal filter structure.

These finite-structured filters restrict the space of filter impulse responses to a finite dimensional subspace of L_2. That is, the optimization is carried out on this subspace rather than on the whole of L_2. The subspace spanned by $\{v_i\}_{i=0}^{n-1}$ (denoted as $\left[\{v_i\}_{i=0}^{n-1}\right]$) is closed and convex, hence, the feasible region $\mathscr{F} \cap \left[\{v_i\}_{i=0}^{n-1}\right]$ is also closed and convex. Strict convexity of the cost functional ensures a unique optimum solution if an interior point exists. For a norm cost, it follows from Theorem 3: that the sub-optimum CCR filters for the finite filter structures converge to the optimum CCR filter as the filter structure grows. For a general convex cost, the sub-optimum costs converge to the optimum cost.

3.8.2. Hybrid filters

So far we have considered filters with impulse responses in l_2. Now we narrow down the discrete-time filters to practical filters that have a finite structure.

Let $\{v_i\}_{i=0}^{\infty}$ be a total sequence in l_2, and let the impulse response of finite filter structures be described by

$$u = K\boldsymbol{a} = \sum_{k=0}^{n-1} a_k v_k \,,$$

where $K : \boldsymbol{R}^n \to l_2$ is a bounded linear operator.

Such filters can be realized by FIR filter or discrete-time Laguerre networks where, incidentally, the corresponding v_k's are also orthonormal. The set of feasible filter coefficients is given by

$$\left\{\boldsymbol{a} \in \boldsymbol{R}^n : \varepsilon^- \leq \Xi_s K\boldsymbol{a} \leq \varepsilon^+\right\}. \tag{28}$$

Assuming that the noise samples are statistically independent, the average output noise power can be expressed in terms of $\boldsymbol{a}$ yielding the cost functional

$$f(K\boldsymbol{a}) = \boldsymbol{a}^T L\boldsymbol{a} \,, \tag{29}$$

where

$$L = \left(N_0/\tau\right)\int_{-\infty}^{\infty} \boldsymbol{w}(\zeta)\boldsymbol{w}^T(\zeta)d\zeta \,, \tag{30}$$

$$\boldsymbol{w}(t) = \left[w_0(t), \ldots, w_{n-1}(t)\right]^T, \quad w_k(t) = \sum_{j=0}^{\infty}\Lambda(t - j\tau)v_k(j). \tag{31}$$

If the noise samples are independent, the cost functional $\boldsymbol{a}^T L\boldsymbol{a}$ is a strictly convex function of $\boldsymbol{a}$.

4. Numerical examples

In this section, we present three applications of EC filter design to illustrate the effectiveness of the techniques presented in the previous sections. The reader is referred to [9, 15] for many more examples.

4.1. EC with uncertain input

Consider the compression of a 13-bit Barker coded signal shown in Fig. 12.

For this example, we use a 27-tap FIR filter and a Bessel post-filter of 3th order with cut-off frequency $\omega_c = 2\pi/\beta$. To obtain the approximate solution, we have constrained the output at every $t_i = i\beta/8$. Figures 13 and 14 show, respectively, the filter's responses to the nominal input and signals which were randomly perturbed about the nominal input but still fit inside the input mask. Observe that these responses stay within the boundary of the output envelope. The noise gain of the ECUI filter is 0.027.

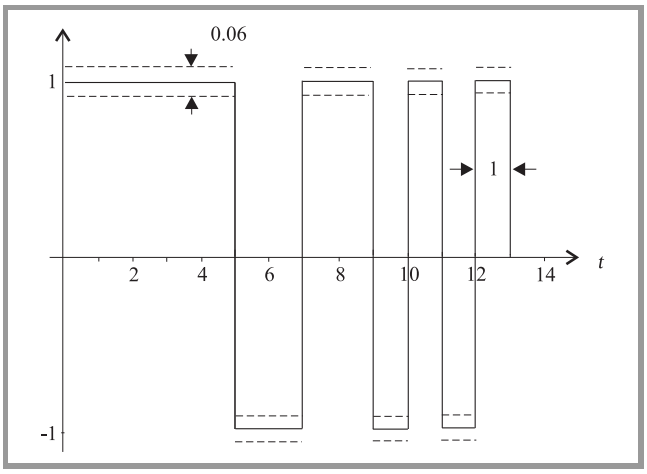


Fig. 12. 13-bit Barker code with input uncertainty.

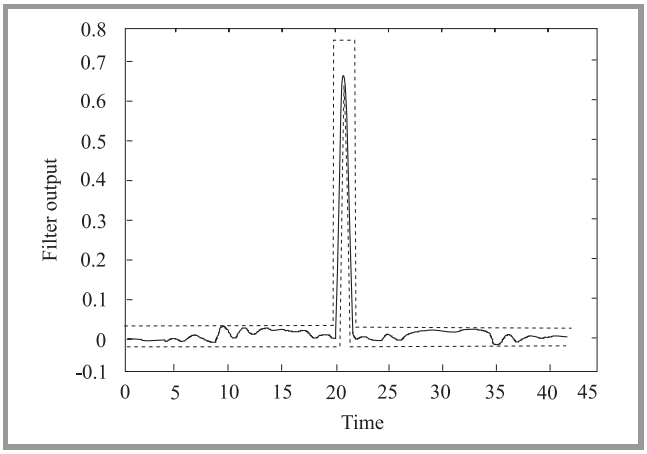


Fig. 13. Response of filter to nominal Barker coded input.

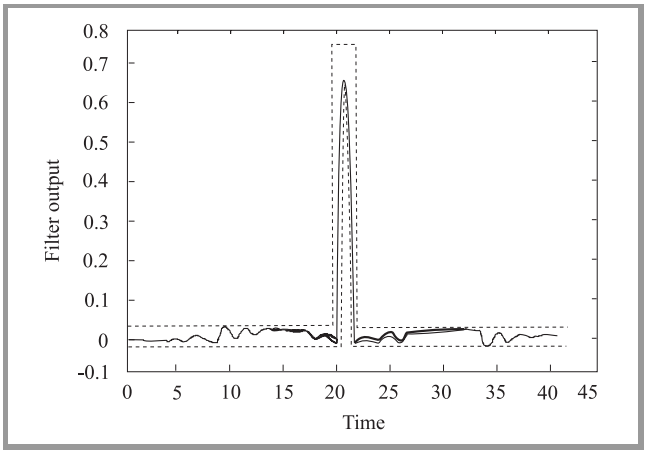


Fig. 14. Response of filter to perturbed Barker code.

4.2. Filter implementation uncertainty

The FALC54 is a Siemens IC for communications used in framing and line interface for PCM30 and PCM24. It

provides a very flexible way to create a custom waveform. Each pulse is partitioned into four subpulses of equal support as shown in Fig. 15. Hence, a pulse u generated from a FALC54 can be modeled as

$$u(t) = \sum_{k=0}^{n-1} x_k \Pi(t - kT/n) \,,$$

where $n = 4$, Π denotes a rectangular pulse with support $[0, T/n]$ and x_k denotes the level of the kth subpulse. The shape of the transmit pulse is thus characterized by the vector of parameters $\boldsymbol{x} = [x_0, x_1, \ldots, x_{n-1}]^T$. We are concerned with the problem of determining the levels of the transmit pulse to satisfy constraints which have arisen from the need to conform to standards such as ANSI, CCITT [3, 22].

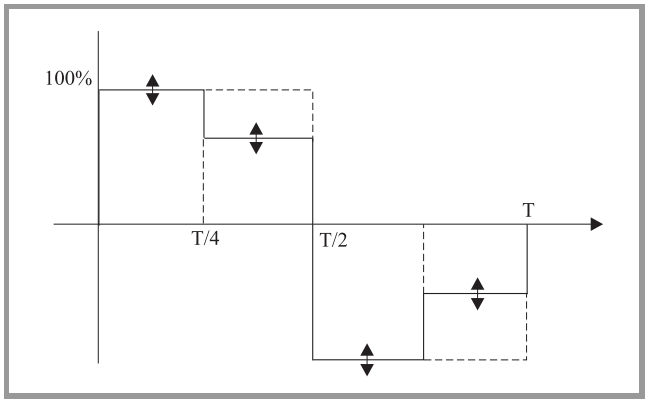


Fig. 15. Partition of pulse support.

Consider the transmission of the T1 pulse over an RG59B/U coaxial cable channel with an attenuation that follows an approximate $\sqrt{f}$ law. For a coaxial cable of length l, its frequency response can be modeled as $H(j\omega) = e^{-Al\sqrt{j\omega}}$. Our design is for a length of a coaxial cable which has an attenuation of 4.5 dB at the DS1 rate $f_0 = 772$ kHz, (i.e. $\omega_0 = 4.8506 \cdot 10^6$ rad/s) and $T = 648 \cdot 10^{-9}$ s. We consider implementation errors on the transmit pulse shape that are bounded by $\boldsymbol{w} = 0.052[1, \ldots, 1]^T$, i.e. $\delta = 0.052$.

Figure 16 shows the received signal at the end of the cable for a optimal transmit pulse design that does not take into account implementation errors bounded by $\delta = 0.052$. Note that for some implementation errors the received pulses violate the prescribed mask.

Figure 17 shows the received signal at the end of the cable for an optimal transmit pulse design that takes into account implementation errors bounded by $\delta = 0.052$ and uses the optimization problem formulated in Section 4.2. Note that the received pulses never violate the prescribed mask for any implementation error that satisfy the assumed bound. The robustness against implementation errors has been obtained at a very low cost of a 9% increase in the transmit pulse energy.

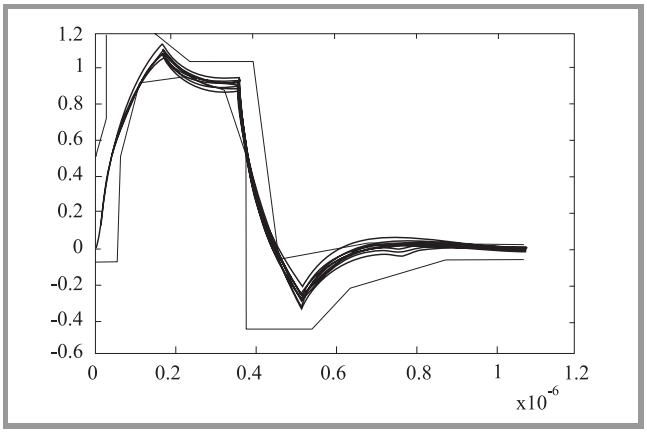


Fig. 16. Received pulse without implementation uncertainty constraints.

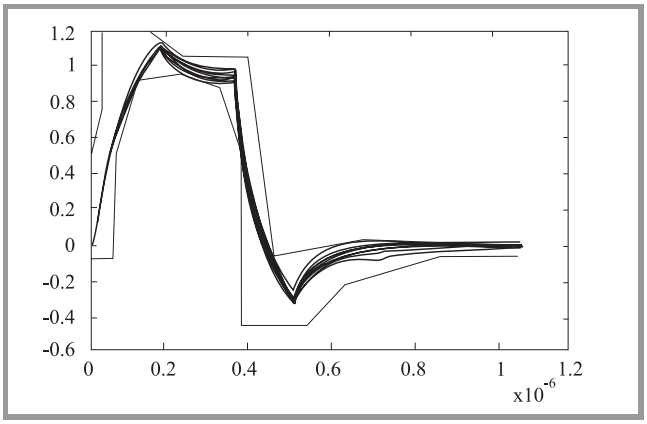


Fig. 17. Received pulse with implementation uncertainty constraints.

4.3. Constraint robustness with Laguerre filter

Let us again consider the design of an equalization filter for a digital transmission channel consisting of a coaxial cable on which data is transmitted according to the DX3 standard [2]. The design objective is to find an equalizer which takes the impulse response of a coaxial cable with a loss of 30 dB at the Baud frequency and produces an output which lies within the DSX3 pulse template. For computational purpose, we consider the time domain constraints at a finite number of discrete points rather than the entire continuum.

In our numerical studies, we use 1024 points over an interval of $[0, 32\beta]$. Using the Laguerre filter with 14 coefficients and $p = 12$, we first solve the optimal EC filtering problem without robustness considerations. The optimal noise gain is

$$\|u^0\|^2 = \|\boldsymbol{a}^0\|^2 = 54.2008 \,.$$

The output mask, input signal and the output are shown in Fig. 18.

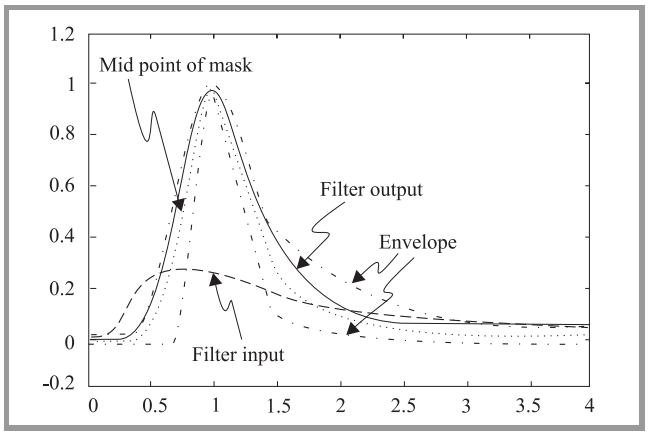


Fig. 18. Laguerre filter – EC approach.

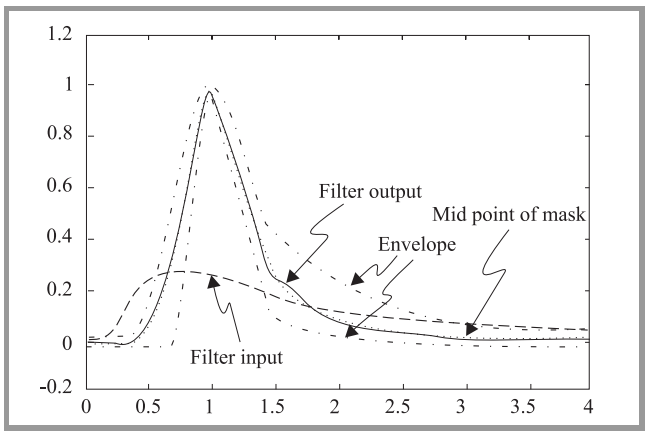


Fig. 19. Laguerre filter – EC approach with robustness constraints.

To achieve maximum constraint robustness margin, we solve the robust EC filtering problem with $\delta = 1.5$ and

$$\beta(t) = \begin{cases} 3\varepsilon(t), & |\varepsilon(t) \geq 0.05 \\ 0.035, & |\varepsilon(t) \leq 0.03 \\ \varepsilon(t), & \text{elsewhere} \end{cases} .$$

The corresponding signals are shown in Fig. 19. From an examination of Fig. 19, we see that the output is very close to the center of the output mask. This increased margin has been achieved by allowing an increase of 50% in the noise gain of the robust EC filter relative to the nosie gain of the non-robust EC filter ($\delta = 1.5$).

5. Conclusion

The paper has reviewed key ideas in envelope constrained filter design. We have shown that the design problem is motivated by practical engineering requirements. A number of filter design examples have been described. These examples include filters for communications systems and pulse compression applicable to sonar and radar ranging systems. The effectiveness of the robust EC filter design approach presented in the paper is evident in these examples.

References

[1] Bell Communications, Technical Reference, TR-TSY-000499, Dec. 1988, issue 2, pp. 9–17.

[2] CCITT, "Physical/Electrical Characteristics of Hierarchical Digital Interfaces", G.703, Fascicle III, 1984.

[3] R. A. Nobakht and M. R. Civanlar, "Optimal pulse shape design for digital communication systems by projections onto convex sets", *IEEE Trans. Commun.*, vol. 43, pp. 2874–2877, 1995.

[4] J. W. Lechleider, "A new interpolation theorem with application to pulse transmission", *IEEE Trans. Commun.*, vol. 39, no. 10, pp. 1438–1444, 1991.

[5] R. J. McAulay and J. R. Johnson, "Optimal mismatched filter design for radar ranging detection and resolution", *IEEE Trans. Inform. Theory*, vol. 17, pp. 696–701, 1971.

[6] A. W. Rihaczek, *Principles of High-Resolution Radar*. New York: Macro-Hill, 1969.

[7] A. Seyler and J. Potter, "Waveform testing of television transmission facilities", in *Proc. IRE*, Australia, July 1960, pp. 470–478.

[8] Z. L. Budrikis, "Visual fidelity criterion and modelling", *Proc. IEEE*, vol. 60, no. 7, pp. 771–779, 1972.

[9] B. Vo, "Optimum envelope constrained filters", Ph.D. thesis, ATRI, Curtin University of Technology, Western Australia, Jan. 1997.

[10] T. E. Fortmann and M. Athans, "Optimal filter design subject to output sidelobe constraints: theoretical considerations", *JOTA*, vol. 14, no. 2, pp. 179–197, 1974.

[11] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.

[12] E. Polak and D. Q. Mayne, "An algorithm for optimization problems with functional inequality constraints", *IEEE Trans. Automat. Contr.*, vol. 21, pp. 184–193, 1976.

[13] G. Gonzaga, E. Polak, and R. Trahan, "An improved algorithm for optimization problems with functional inequality constraints", *IEEE Trans. Automat. Contr.*, vol. 25, pp. 49–54, 1980.

[14] B. Vo, A. Cantoni, and K. L. Teo, "Computational methods for a class of functional inequality constrained optimization problem", in *World Scientific Series in Applicable Analysis – Recent Trends in Optimization Theory and Applications*, R. P. Agarwal, Ed. Singapore: World Scientific Publishing, 1995, vol. 5, pp. 447–465.

[15] R. J. Evans, T. E. Fortmann, and A. Cantoni, "Envelope-constrained filter". Part I. "Theory and applications", *IEEE Trans. Inform. Theory*, vol. 23, pp. 421–434, 1977.

[16] R. J. Evans, A. Cantoni, and T. E. Fortmann, "Envelope-constrained filter". Part II. "Adaptive structures", *IEEE Trans. Inform. Theory*, vol. 23, pp. 435–444, 1977.

[17] W. X. Zheng, A. Cantoni, B. Vo, and K. L. Teo, "Recursive procedures for constrained optimization problems and its application in signal processing", *IEE Proc. – Vision, Image Signal Proc.*, vol. 142, no. 3, pp. 161–168, 1995.

[18] T. E. Fortmann and R. J. Evans, "Optimal filter design subject to output sidelobe constraints: computational algorithm and numerical results", *JOTA*, vol. 14, no. 3, pp. 271–290, 1974.

[19] L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*. New Jersey: Prentice-Hall, 1975.

[20] B. Vo, Z. Zang, A. Cantoni, and K. L. Teo, "Continuous-time envelope constrained filter design via orthonormal filters", *IEE Proc. – Vision, Image Signal Proc.*, vol. 142, no. 6, pp. 389–394, 1995.

[21] B. Vo, T. Ho, A. Cantoni, and V. Sreeram, "FIR filters in continuous-time envelope constrained filter design", in *Proc. IEEE ICASSP*, Munich, Germany, 1997, vol. 3, pp. 1905–1908.

[22] R. J. Evans, A. Cantoni, and K. M. Ahmed, "Envelope-constrained filters with uncertain input", *Circ. Syst. Signal Proc.*, vol. 2, no. 2, pp. 131–154, 1983.

[23] K. L. Teo, A. Cantoni, and X. G. Lin, "A new approach to the optimization of envelope constrained filters with uncertain inputs", *IEEE Trans. Signal Proc.*, vol. 42, no. 2, pp. 426–428, 1994.

[24] W. X. Zheng, A. Cantoni, and K. L. Teo, "Robust design of envelope constrained filters in the presence of input uncertainty", *IEEE Trans. Signal Proc.*, vol. 44, no. 8, pp. 1872–1878, 1996.

[25] W. Zheng, A. Cantoni, and K. L. Teo, "The sensitivity of envelope-constrained filters with uncertain input", *IEEE Trans. Circ. Syst. – I,* vol. 42, no. 9, 1995.

[26] B. A. Francis, "A course in H_∞ control theory", in *Lecture Notres in Control and Information Sciences*, M. Thomas and A. Wyner, Eds. Berlin: Springer Verlag, 1987.

[27] B. Vo and A. Cantoni, "Continuous-time enevlope constrained filter design with input uncertainty", *IEEE Trans. Circ. Syst. – I*, vol. 47, no. 10, pp. 1445–1454, 1995.

Antonio Cantoni (FIEEE, FTSE) was born in Soliera, Italy, on 30 October, 1946. He received the B.E. (first class honours) and Ph.D. degrees from the University of Western Australia, Nedlands, in 1968 and 1972, respectively. He was a Lecturer in Computer Science at the Australian National University, Canberra, in 1972. He joined the Department of Electrical and Electronic Engineering at the University of Newcastle, Shortland, NSW, Australia in 1973, where he held the Chair of Computer Engineering until 1986. In 1987, he joined QPSX Communications Ltd, Perth, Western Australia, as Director of the Digital and Computer Systems Design Section for the development of the DQDB Metropolitan Area Network. From 1987 to 1990 he was also a visiting Professor in the Department of Electrical and Electronic Engineering at the University of Western Australia. From 1992 to 1997 he was the Director of the Australian Telecommunications Research Institute and Professor of Telecommunications at Curtin University of Technology, Perth, Western Australia. During this period he was also the director of the Cooperative Research Centre for Broadband Telecommunications and Networking. He is currently Chief Technology Officer with Atmosphere Networks Inc and Professor of Telecommunications at the University of Western Australia. He is interested in adaptive signal processing, electronic system design, and networking and regularly acts as a consultant to industry in these areas. Dr Cantoni is a Fellow of the IEEE and a Fellow of the Australian Academy of Technological Sciences and Engineering. He has been an associate editor of the IEEE Transactions on Signal Processing.
e-mail: cantoni@ec.uwa.edu.au
University of Western Australia
Australia

Ba-Ngu Vo (MIEEE) was born in Saigon Vietnam. He received the B.Sc./B.E. degree (with first class honours) in 1994 from the University of Western Australia and his Ph.D. degree in 1997 from the Australian Telecommunications Research Institute (ATRI), Curtin University of Technology. He is currently a research fellow at the Department of Electrical and Electronic Engineering, the University of Melbourne. His research interests lie in the application of optimization techniques to signal processing, semi-infinite programming, filter design, tracking and data fusion.
e-mail: bv@ce.mu.oz.au
The University of Melbourne
Australia

Kok Lay Teo received the Ph.D. degree in electrical engineering from the University of Ottawa, Ottawa, ON, Canada in 1974. He was with the Department of Applied Mathematics, the University of New South Wales, Sydney, Australia, from 1974 to 1985, as a Lecturer and subsequently as a Senior Lecturer. He was an Associate Professor in the Department of Industrial and Systems Engineering, National University of Singapore from 1985 to 1987. He was an Associate Professor in the Department of Mathematics, The University of Western Australia, Nedlands, Australia, from 1988 to 1996. He was Professor of Applied Mathematics at Curtin University of Technology, Perth, Australia, from 1996 to 1998. He is currently the Chair Professor of Applied Mathematics and Head of the Department of Applied Mathematics at The Hong Kong Polytechnic University, Hong Kong. He has delivered 6 keynote lectures, 4 fully funded invited lectures, and published 4 books and numerous journal and conference papers. The software package, MISER3.2, for solving general constrained optimal control problems was developed by the research team under his leadership. His research interests include both the theoretical and practical aspects of optimal control and optimization, and their applications, in particular, to signal processing in telecommunications.
e-mail: teo@maths.curtin.edu.au
The Hong Kong Politechnic University
Hong Kong

Paper

Source enhanced linear prediction of speech incorporating simultaneously masked spectral weighting

Jason Lukasiak and Ian S. Burnett

Abstract — Linear prediction is the cornerstone of most modern speech compression algorithms. This paper proposes modifying the calculation of the linear predictor coefficients to incorporate a weighting function based on the simultaneous masking property of the ear. The resultant prediction filter better models the perceptual characteristics of the source and results in the removal of more perceptually important information from the input speech signal than a standard LP filter. When employed in a low rate speech codec the net effect is an improvement in subjective quality, with no increase in transmission rate and only a modest increase in computational complexity.

Keywords — linear prediction, psychoacoustics, masking, LPC.

1. Introduction

Linear prediction (LP) forms an integral part of almost all modern day speech coding or speech compression algorithms. The primary reason for this popularity is that linear prediction provides a relatively simple and well founded technique for removing the redundancy from a speech signal, thus aiding in compression or bit rate reduction. Linear prediction determines and removes redundancy by removing the short term correlations of the input signal.

Whilst linear prediction is widely used in speech coding it was not originally developed specifically for speech coding but rather for the more general field of signal processing. The result of this is that the linear predictor used for speech coding does not exploit many of the well known perceptual properties of human hearing. These perceptual properties include the nonlinear frequency response of the ear and simultaneous masking, amongst others and are well defined in texts such as [1]. Previous authors [2 – 4] have incorporated some perceptual properties into the calculation of the linear predictive filter. These authors have reported good results, primarily by warping the frequency axis to simulate the nonlinear frequency response of the ear prior to calculating the filter parameters. Hermansky [4] also included equal loudness perception and the intensity-loudness power law into the calculation of the filter. Whilst these authors reported good results none of them attempted to incorporate simultaneous masking into the filter calculation.

Simultaneous masking occurs in the frequency domain when a high amplitude sound causes adjacent lower amplitude sounds to become inaudible. This property has been widely used in many audio coding techniques, such as MPEG4 [5], as a tool to determine the optimal quantization step size required to code the input and thus allows perceptually transparent compression of the audio signal. However, the use of simultaneous masking in speech coding algorithms has been very limited due to the increased computational demand required to perform quantization of the signal in the frequency domain.

This paper proposes a method for modifying the calculation of the linear prediction coefficients (LPC) to better model the characteristics of the source. This is achieved by incorporating a weighting function based on the simultaneous masking property of the ear into the calculation of the LPC. This approach fits the linear predictive spectrum only to the unmasked samples of the input spectrum. The motivation for this technique is to ensure no complexity is wasted modeling the masked regions, thus allowing the unmasked regions to be better represented. This allows the filter to remove more perceptually important information from the signal than the standard technique, with the resultant residual signal consisting of less perceptually important information. This characteristic allows the subjective quality of the synthesized speech to be improved for a given residual quantization scheme. This paper presents results confirming this characteristic using objective measures and subjective listening tests.

The paper is organized as follows. In Section 2 an overview of linear prediction and human auditory perception is detailed. The new linear prediction method is given in detail in Section 3. Objective and subjective results are provided in Section 4. Finally, the major points are summarized in Section 5.

2. Background

2.1. Linear prediction analysis

The use of a linear predictor in speech coding relies upon the fact that speech can be modeled as the output of a time varying linear system [6]. The development of this model is linked to the use of lossless acoustic tubes to represent the speech production process and is detailed in [6]. Figure 1 represents a simplified representation of this model.

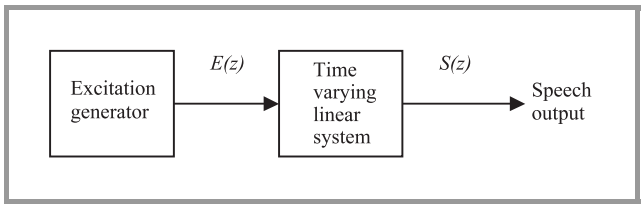


Fig. 1. Source-system model of speech production.

The transfer function representing the linear system in Fig. 1 can be described using an all pole (autoregressive (AR)) system as:

$$H(z) = \frac{S(z)}{E(z)} = \frac{G}{1 - \sum_{k=1}^{p} a_k z^{-k}}, \qquad (1)$$

where p is the order of the filter, G is the gain of the system and the a_k are the predictor coefficients.

The popularity of the AR model stems from the fact that if a sufficient order is used the filter can accurately model the speech production system and also the filter parameters can be calculated in a straightforward and efficient manner [6].

To allow linear prediction the input waveform must be a stationary random process. However, as speech is a nonstationary random process some modification is required. It has been determined that a speech signal is stationary over a period of approximately $20-30$ ms [7]. Thus to utilize linear prediction in speech coding it is necessary to divide the input speech into frames of approximately 20 ms length and update the linear prediction coefficients for each of these frames. A number of methods to solve for the predictor coefficients to achieve a minimum mean square error for a given frame have been developed [6]. The most popular of these is the autocorrelation method and this is the method used in this paper.

The mean square error (MSE) solution for the standard LPC's (a_k) can be reduced using the autocorrelation method [8], to:

$$R(l) = \sum_{k=1}^{p} a_k R(l-k), \quad l = 1 \ldots p \qquad (2)$$

where $R()$ is the autocorrelation function of the input speech frame. The recursive Levinson-Durbin [9] algorithm is then used to solve for the filter coefficients a_k.

2.2. Overview of human auditory perception

The human auditory system is a highly complex system. Sounds presented to the ear are not all perceived equally but are governed by a number of nonlinear operations. Humans can hear sounds in the range of approximately 50 Hz $-$ 16 kHz [1], however, these frequencies are not all perceived with equal sensitivity. This phenomenon leads to a set of curves called equal loudness curves [1] which indicate the perceived loudness of a fixed amplitude tone, as the frequency of the tone is varied. Directly related to the equal loudness curves is the threshold of hearing curve. This curve represents the minimum amplitude that is audible for a given frequency.

The frequency scale of the human ear also acts as a set of overlapping bandpass filters. These filters are called critical band filters and the pass band of each individual band pass filters is termed a critical band. The nature of each critical band is that all frequencies within a band are perceived equally by the ear [1]. The critical bands are not of equal width but increase in bandwidth as their center frequency increases. This results in better frequency resolution at lower frequencies. Scharf [10] proposed that the critical bands could be adequately represented by a set of non overlapped rectangular filters. This greatly reduces the complexity of critical band analysis.

Masking occurs when a loud sound causes other softer sounds to become inaudible. Two types of masking can occur, namely simultaneous and temporal masking. Simultaneous masking occurs when a low intensity but audible sound is made inaudible by a higher intensity adjacent sound occurring at a simultaneous moment in time. Temporal masking occurs when a loud tone causes softer tones occurring before and after the tone to become inaudible [1].

The masking and equal loudness phenomena have been researched extensively and have led to the development of psychoacoustic models that allow the auditory system to be accurately modeled. Detailed descriptions of these models can be found in many well recognized texts such as [1].

Incorporating the perceptual characteristics of the auditory system into audio and speech coding allows coders to operate more efficiently at reduced bit rates by distributing the available bits according to the perceptual nature of the signal. This allows the coder to reproduce a perceptually unaltered signal at a greatly reduced bit rate when compared with coders that ignore the perceptual characteristics. An example of this principle is the MPEG4 [5] audio coding standard. Perceptual modeling provides the crux of this coder and allows the coder to produce high quality audio signals at a relatively low bit rate.

3. Linear prediction incorporating simultaneously masked spectral weighting (SMWLPC)

3.1. Motivation

The motivation for this technique was to allow a simple and computationally efficient means of better exploiting the perceptual characteristics of human hearing in low rate LP based speech codecs. Traditionally, perceptual distortion is only exploited in low rate LP speech coding by using a noise shaping weighting filter [11] when coding the resid-

ual signal. This filter is used to weight the error signal when searching for the optimal excitation signal to represent the LP residual. This weighting filter de-emphasises the frequency regions corresponding to the formants of the input speech. This de-emphasis exploits the masking characteristic of the ear in that larger errors are imperceptible in louder sections of the input speech than in quieter sections. Whilst the use of this weighting filter has produced good results [12] and been widely accepted, it uses only a basic model of the perceptual characteristics of the ear. Authors such as Sen [13] and Burnett [14] have reported improved performance by employing more sophisticated perceptual models when searching for the excitation signal that minimizes the perceptually weighted MSE to the LP residual. The improved results reported by [13] and [14] have been at the expense of a large increase in computational complexity. This increase in complexity is due to the fact that each perspective excitation signal tested must be transformed to the frequency domain and the error signal then multiplied with the respective perceptual model.

SMWLPC attempts to exploit more sophisticated perceptual models than the filter proposed in [11]. Incorporating these models into the LP filter allows SMWLPC to remove more perceptually important information from the input signal than standard LPC thus resulting in a residual signal that contains less perceptually important information. By exploiting the perceptual models upfront in the LP filter, computational complexity is dramatically reduced when compared to methods where the complex perceptual models are used when quantizing the LP residual [13]. This reduction is due to the fact that only a single transform and weighting multiplication per frame of input speech is required. Also by incorporating the complex perceptual models into the LPC, SMWLPC can be easily adapted to any existing LP based speech coding algorithm by simply replacing the standard LP filter.

The method selected limits the increase in computational complexity over standard LP filtering by maintaining the use of traditional recursive solutions in calculation of the LPC's. This also maintains a stable autoregressive structure that can be directly employed in any LP based speech codec. Standard linear prediction minimizes the error equally across the entire frequency spectrum of the input speech. This approach fails to exploit many of the well known perceptual properties of hearing. The SMWLPC technique employs a method of incorporating simultaneous masking into the calculation of the linear prediction coefficients. This allows the error to be minimized only in the sections of the input spectrum that are unmasked. This is achieved by first determining which frequencies in the input signal are simultaneously masked and then ignoring them in the calculation of the LPC.

3.2. Method

A block diagram of the SMWLPC method is shown in Fig. 2. Initially the power spectrum (frequency domain) of the input speech is calculated via a fast Fourier transform (FFT). A masking threshold function is then calculated for each discrete frequency. The calculation of this function is detailed in Section 3.3. The masked input frequencies are then determined. This is achieved by comparing the power spectrum of each discrete frequency to the masking threshold for that frequency. If the power spectrum is less than the masking threshold or the threshold of hearing, the frequency is deemed masked. A modified power spectrum is then produced by taking those frequencies deemed masked and zeroing their value. This method is equivalent to generating a spectral weighting function whose values are unity for unmasked frequencies and zero for masked frequencies or frequencies whose power is below the threshold of hearing and then multiplying the input spectrum by this weighting function. The result is a power spectrum that contains only unmasked information. Recognizing that the autocorrelation of a discrete stochastic signal is the inverse discrete Fourier transform (IDFT) of the power spectrum, the perceptually altered power spectrum is transformed to the autocorrelation function of the unmasked speech. A perceptually altered linear predictor can then be easily calculated using the well known Levinson-Durbin recursion [9].

3.3. The masking threshold function

The psychoacoustic model used to calculate the masking threshold function is based on that proposed in [15] with the parameters modified to optimize the performance of the SMWLPC. A block diagram of the method is shown in Fig. 3.

The input power spectrum is segmented into N non overlapped critical bands. Where N represents the number of critical bands that exist within the bandwidth of the input signal. For narrowband speech the input bandwidth is approximately 4 kHz and the number of critical bands is 18. The critical bands are described in Section 2.2 and are given in detail in [10]. The power spectrum lines within each critical band are summed together, this gives an energy estimate for each critical band. The combination of the N energy estimates is called the band energy waveform as referred to in Fig. 3.

To simulate the masking effect between critical bands, the band energy waveform is provided as the input to an inter-band masking calculator. The inter-band masking calculator convolves the band energy waveform with a spreading function (shown in Fig. 4) to produce a spread band energy waveform. The spreading function shown in Fig. 4 is identical to that given in [1] and has been derived from exhaustive psychoacoustic testing.

The spread band energy waveform is then used to determine an initial masking threshold function (IMTF) according to the following formula:

$$IMTF(i) = Energy(i) - O(i), \tag{3}$$

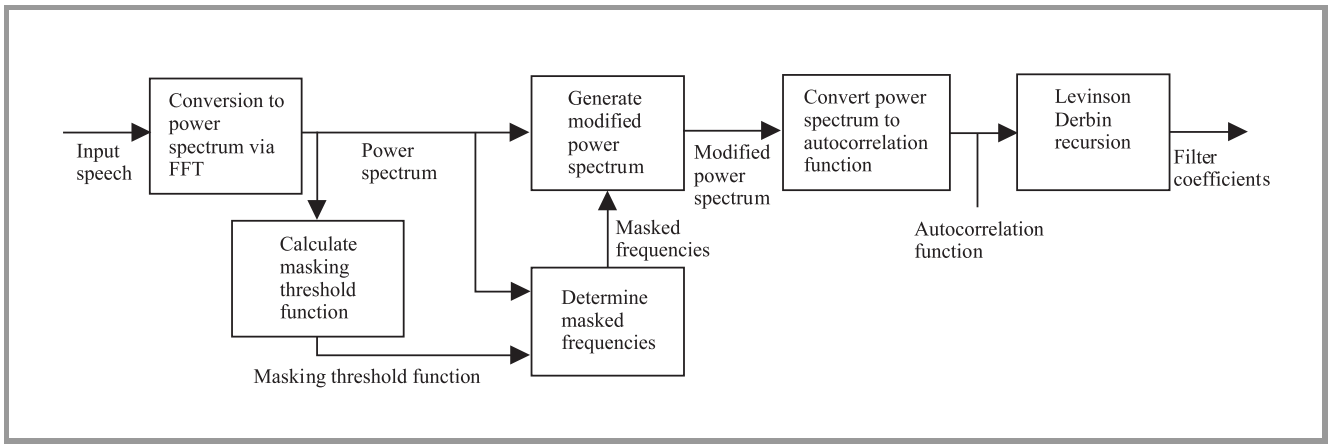


Fig. 2. Functional block diagram of SMWLPC.

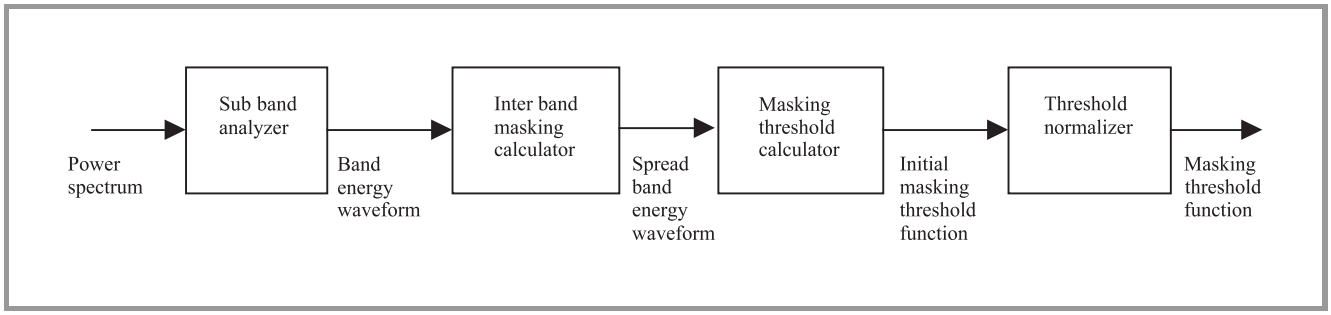


Fig. 3. Block diagram of the masking threshold calculation.

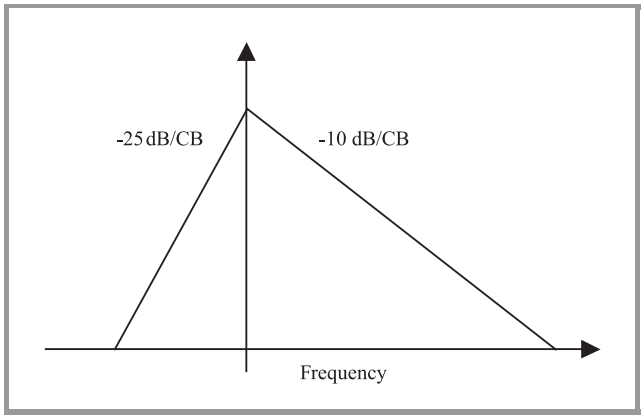


Fig. 4. Inter-band spreading function.

where $Energy(i)$ represents the total energy of the ith band of the spread band energy waveform measured in decibels; $O(i)$ is given by:

$$\begin{aligned} O(i) &= 50, && \alpha < 0.2 \\ O(i) &= \alpha(\beta + i) + (1-\alpha)\gamma, && \alpha \geq 0.2 \end{aligned} \tag{3.1}$$

where:

$$\alpha = \min\left(\frac{SFM}{SFM_{\max}}, 1\right), \tag{3.2}$$

$$SFM = 10 \log \frac{G_m}{A_m}, \tag{3.3}$$

$SFM_{\max}$ is an empirically determined value; G_m is the geometric mean of the power spectrum; A_m is the arithmetic mean of the power spectrum; β and γ are empirically determined constant values that represent the tone masking noise and noise masking tone thresholds respectively.

The value of $SFM_{\max}$ suggested in [15] was -60 dB however, upon testing with a pure sine wave at 1 kHz the required $SFM_{\max}$ to give an alpha value of 1 was determined to be -40 dB. In Eq. (3.1) β and γ are set to 14.5 and 7 respectively; α is a measure of the flatness of the power spectrum, a value of 1 indicates a purely tonal signal and 0 represents pure noise. Equation (3.1) utilizes α to ensure that the correct mix of noise and tone thresholds is selected. The value for $O(i)$ in Eq. (3.1) differs greatly from that suggest in [15] where the definition is given as:

$$O(i) = \alpha(\beta + i) + (1-\alpha)\gamma \quad \text{for all } \alpha. \tag{4}$$

Setting Eq. (3.1) to a very large constant value for a very noise like signal ($\alpha < 0.2$) ensures that for this type of input the IMTF is made small and designates virtually the entire spectrum unmasked. This overcomes the situation where Eq. (4) designates virtually the entire spectrum masked for such a signal thus leaving too few samples to successfully generate the filter coefficients. This characteristic was reported to cause distortions in the reconstructed speech in [16] and the modification overcomes this problem. Also in Eq. (4) $\gamma = 5.5$, contrasting with the increased $\gamma = 7$ in Eq. (3.1). This modification enhances the performance of

SMWLPC and was determined empirically through informal listening tests.

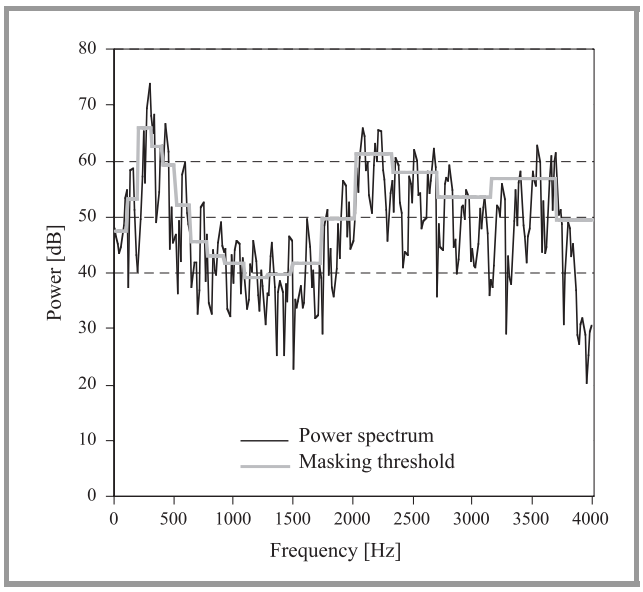


Fig. 5. Example of the masking threshold function.

The IMTF is then adjusted by the threshold normalizer to account for misestimation of the $Energy(i)$ values resulting from the shape of the spreading function. This results in a masking threshold function an example of which (with the corresponding power spectrum) is shown in Fig. 5.

3.4. Mathematical analysis of SMWLPC

SMWLPC is now analyzed mathematically to explore and contrast the differences between this approach and standard LPC.

The MSE solution for the standard linear predictive coefficients using the autocorrelation method was given in (2). As the input frame of speech is assumed to be a stationary random process the autocorrelation values $(R(n))$ can be computed via an inverse discrete Fourier transform of the power spectral density $P(k)$ [17]:

$$R(n) = \frac{1}{N}\sum_{k=0}^{N-1} P(k)e^{jwkn/N} \quad n = 0 \ldots N-1. \tag{5}$$

If the calculation of $R(n)$ in (5) is modified to only operate on the perceptually important (unmasked) values of k then the autocorrelation becomes:

$$R(n) = \frac{1}{L}\sum_{unmasked\ \ l} P(l)e^{jwln/N} \quad n = 0 \ldots N-1, \tag{6}$$

where L represents the number of unmasked frequency bands of N.

Substituting the autocorrelation sequence (6) into (2) gives:

$$\frac{1}{L}\sum_{unmasked\ \ l} P(l)e^{jwln/N} =$$

$$= \sum_{k=1}^{p} a_k \left(\frac{1}{L}\sum_{unmasked\ \ l} P(l)e^{jwl(n-k)/N}\right), \ n = 1 \ldots p. \tag{7}$$

It is clear that the above equations solve the mean square solution for a_p using only the unmasked values of k. Also as $\frac{1}{L}$ is a common factor it can be removed from the equations. This results in each summation term being equal to only the sum of the unmasked values of $P(K)$ multiplied by the respective harmonic component, which is identical in value to the sum over all k with the masked values of $P(K)$ set to zero.

The above analysis demonstrates that SMWLPC fits only to unmasked regions and simply ignores the masked regions in its calculation of the LP coefficients. The fact that only the unmasked regions are modeled allows SMWLPC to achieve a better fit to these regions as complexity is not wasted attempting to model masked regions.

An alternate approach to examining the effect of the SMWLPC is to view the predictor error in the frequency domain. This can be expressed as [6]:

$$E = \frac{G^2}{2\pi}\int_{-\pi}^{\pi} \frac{|S(e^{jw})|^2}{|H(e^{jw})|^2}\, dw. \tag{8}$$

Equation (8) shows that minimizing E is equivalent to minimizing the ratio of the input energy spectrum $(S(e^{jw}))$ to the squared magnitude of the frequency response $(H(e^{jw}))$ of the model. It can be seen that zeroing the power spectrum (numerator of equation) at any particular frequency, causes the difference between the model and the spectrum at that frequency to contribute nothing to the integral of the ratio over the entire spectrum. The result is that the zeroed (masked) regions have no effect in calculating the linear predictive coefficients.

3.5. Computational complexity

The computational complexity of SMWLPC is increased when compared to the standard LPC. However, this includes calculation of the psychoacoustic model parameters which remain available for other coding tasks such as quantization. In standard LPC, calculation of the autocorrelation requires $(p+1)N_w$ operations [6], where p is filter order and N_w is the window size. The SMWLPC uses an FFT and requires $N_f \log_2 N_f$ multiplications plus $N_f/2$ comparisons to calculate the autocorrelation function, where N_f is the FFT length used. The SMWLPC also requires approximately $2N_f + 700$ operations in calculation of the psychoacoustic parameters. Both methods require approximately p^2 operations to solve the matrix equations. The configuration in this paper used $N_w = 240$, $p = 10$ and $N_f = 512$. The complexities in this case are SMWLPC = 5892 operations and standard LPC = 2740 operations. The computational demand of SMWLPC can be made approximately equal to that of the standard LPC by using an FFT of length 256. This size transform has little effect on the performance of SMWLPC for 4 kHz band limited speech.

3.6. Data windowing requirements

The psychoacoustic model given in [15] was based on audio signals sampled at 32 kHz or greater. Due to this sampling rate and also the ability to accept large delays in audio coding as they seldom operate in real time, the transform length is set to 2048. Using this length window produces frequency bins that are separated by less than 16 Hz and as the lowest frequency in the audio range is > 50 Hz, spectral leakage between frequency bins [17] has little effect even if a rectangular window were used. To adapt this model for use in narrowband speech coding with a sampling rate of only 8 kHz and a constraint on the maximum delay due to the need to operate in real time, the number of samples in the window is greatly reduced. This short transform length causes the frequency separation between adjacent frequency bins to become almost equal to the lowest pitch value for voiced speech of approximately 50 Hz. This characteristic causes spectral leakage across the frequency bins to have a large effect for low pitch speech. The leakage can act as an initial spreading function for low pitched speech and thus causes the masking threshold generated for this speech to become distorted. The authors have found that if a Hamming window of length 240 samples is used the effects of leakage are minimized and results have shown the masking threshold to be consistent across a range of pitch values. This window length is towards the upper limits used in speech coding but is common in low rate speech coders such as the FS1016 [18] 4.8 kbps CELP coder. If a shorter window length is used the spectral leakage can cause the masking thresholds to become inconsistent across the range of possible pitch values.

4. Experimental results

4.1. Objective results

4.1.1. LPC spectral estimate

The spectra of the linear predictive filter provides a good estimate of the spectra of the input speech. This relationship is clearly evident when examining Eq. (8) which shows that the predictor coefficients are calculated by minimising the ratio of squared error between the speech and filter spectra. This property of linear predictive filtering is widely exploited in harmonic coders to provide a bit rate effective means of transmitting the spectral envelope.

To examine the effect of SMWLPC on the accuracy of the spectral estimate, 10th order LPC and SMWLPC analyses were performed for a number of voiced and unvoiced speech segments. The spectra produced by both methods were then compared to the actual speech spectrum. A typical example of the spectrum produced is shown in Fig. 6. The masked frequencies are indicated by shading. It is clearly evident in Fig. 6 that the SMWLPC spectra is a more accurate representation of the input speech spectra in unmasked formant regions. As can be seen at around 800 Hz in Fig. 6 the increased accuracy often results in the

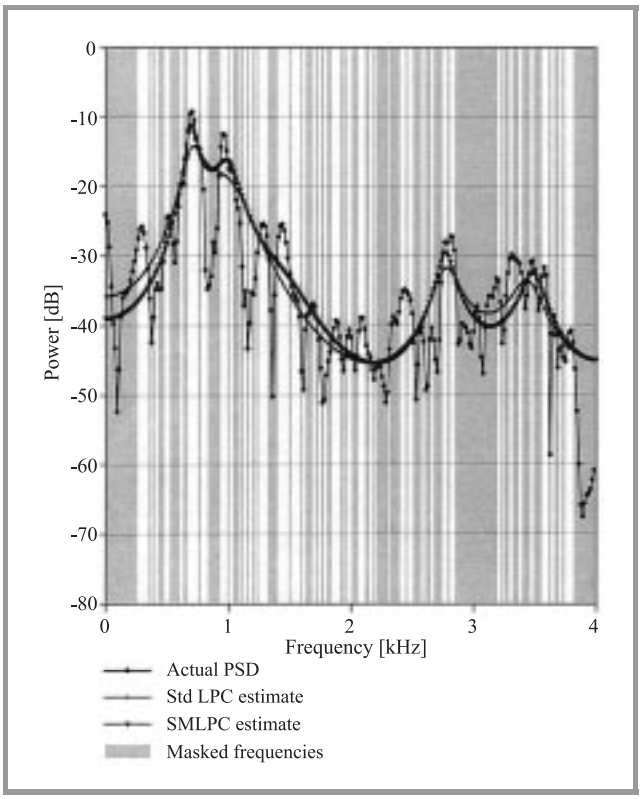


Fig. 6. Comparison of SMWLPC and standard LPC spectral estimates.

SMWLPC modeling 2 distinct formant peaks of the input spectrum whilst the standard LPC produces only a single peak between the two peaks of the input spectrum. This better modeling of the perceptually important formant regions allows SMWLPC to remove more of this perceptually important information from the input speech that a standard LP filter.

To obtain an objective measure for the amount of extra perceptually important information that is removed by SMWLPC the average weighted unmasked residual energy ($WURE$) was calculated using:

$$WURE = \frac{1}{L} \sum_{\text{unmasked } l} w(l) P_r(l)\,, \tag{9}$$

where $P_r()$ is the PSD of the residual signal, $w()$ is a weighting function equal to a 40th order LP spectrum of the input speech normalized to have unity maximum and L is the total number of unmasked spectral lines. The use of the weighting function $w()$ in Eq. (9) places greater emphasis on the perceptually important formant regions of the input spectrum.

10th order LP analysis using both SMWLPC and standard LPC was performed on 10 input sentences (5 male/5 female). Frames of length 200 samples were used in the analysis with a linear predictive Hamming window of 240 samples having an overlap of 20 samples between frames. The $WURE$ of each frame with an α (from Eq. (3.2)) greater than 0.2 was calculated via Eq. (9) and the values averaged

across the entire sentence. An α greater than 0.2 was used as these are the frames for which SMWLPC differs from standard LPC as explained in Section 3.3. The results of the analysis are shown in Table 1.

Table 1
Percentage greater WURE removed by SMWLPC

Sentence number	Speaker gender	SMWLPC % improvement
1	Male	4.7
2	Male	17.52
3	Male	3.95
4	Male	3.55
5	Male	10.5
6	Female	2.87
7	Female	3.84
8	Female	5.82
9	Female	2.51
10	Female	3.13

The results in Table 1 show the average percentage reduction in *WURE* for SMWLPC compared to standard LP for each input sentence. The results demonstrate that more perceptually important information was removed by SMWLPC for each of the input files. The average improvement for all sentences was 5.84%, this represents a significant improvement and indicates that SMWLPC removes significantly more perceptually important information from the input signal than standard LPC.

A typical example of the difference between the weighted residual power spectrums for a standard LP filter and the SMWLPC filter over a typical speech segment is shown in Fig. 7. A positive value indicates that the SMWLPC residual has greater power and a negative signal indicates that the standard LPC residual is of higher power. The masked frequencies are shaded. Figure 7 shows that in ranges of frequency that are largely free of masking or exhibit regular spaced masking (strongly voiced) such as between 0 Hz and 1500 Hz, the SMWLPC residual has lower power than the standard LPC residual. Also in regions that are heavily masked such as between 2700 Hz and 3500 Hz the SMWLPC residual has greater magnitude than the standard LPC residual. These results reinforce the claims that the SMWLPC removes more of the perceptually important unmasked information from the signal than a standard LPC.

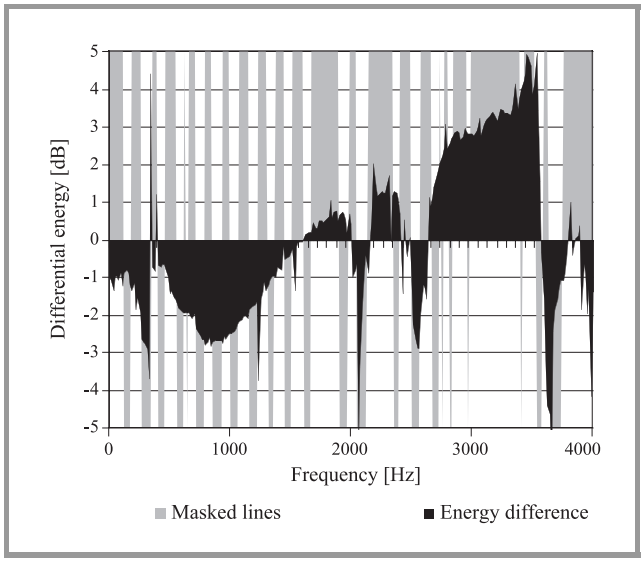


Fig. 7. Difference in weighted residual energy.

4.1.2. Quantization properties

The direct form LP coefficients shown in the calculations of Section 2.1 are susceptible to quantization noise [6]. Due to this characteristic they are rarely used in speech coding [19]. The most popular representation of the LP coefficients are line spectral frequencies (LSF). The LSF's are calculated from the direct form coefficients and their characteristics make them suitable for quantization. These characteristics include monotonically increasing order, strong intra and inter frame correlation and clustering together at formant frequencies [7].

To examine the effect that SMWLPC has on the correlation properties of the LSF's, the inter and intra frame correlation for both standard and SMWLPC LSF parameters were compared. This comparison showed no significant differences in the correlation values between the two methods. To verify this finding in a practical situation a vector linear predictor as proposed in [20] was calculated for both the standard LPC and SMWLPC LSFs respectively. The predictor produced is a square matrix that uses the LSF vector from the previous frame to estimate the LSF vector for the current frame by exploiting both intra and inter frame correlation. The spectral distortion between the predicted vector and the actual vector was calculated for each frame of a test sequence of 1000 frames and was then averaged across all frames. The spectral distortion was calculated via Eq. (10) and the results are shown in Table 2.

$$sd = \sqrt{\frac{1}{\frac{N}{2}} \sum_{k=0}^{\frac{N}{2}-1} \left[10 \log \frac{|S(k)|^2}{|H(k)|^2} \right]^2}, \qquad (10)$$

where N is the FFT length, $S(k)$ is the actual LPC spectrum and $H(k)$ is the predicted LPC spectrum.

Table 2
Average spectral distortion for the predicted LSF vector

	Average spectral distortion [dB]
SMWLPC	2.38
STD LPC	2.31

The results shown in Table 2 indicate that the spectral distortion was virtually identical for both LP methods. The

difference of 0.07 dB is statistically insignificant as the resultant error would then be vector quantized and a final spectral distortion of less than 1 dB is known to produce transparent results for speech coding [21]. Achieving a virtually identical spectral distortion indicates that in practical situations SMWLPC maintains the high inter and intra frame correlation values of standard LPC LSF's and thus are suitable for high compression quantization schemes such as vector linear prediction.

4.2. Subjective listening tests

To test the performance of the SMWLPC in existing speech codecs, a version of the 4.8 kbps FS 1016 CELP coder [18] and a WI [22] coder operating at 2 kbps [16] were modified to use the SMWLPC in place of the standard LPC. The motivation for selecting the CELP and WI coders was to test the performance of SMWLPC in structures that code the LP residual signal in a closed loop and open loop method respectively. As the WI coder uses vector quantization of the LSF parameters the coder was set to operate with non quantized LSF's. This removed the need to retrain the LSF codebook for the SMWLPC and also ensured an unbiased evaluation of SMWLPC's effect on the perceptual content of the residual signal, with no effect from quantization errors of the LPC parameters. This modification was not necessary for the CELP coder as it uses scalar quantization of the LPC parameters which were found to match both the standard LPC and SMWLPC. All other parameters including codebooks were left unaltered.
Each of the coders was used to generate synthesized speech for 10 input speech sentences (5 male, 5 female) from the TIMIT database using both the standard LPC and SMWLPC. Subjective forced A/B comparison testing comprising 20 untrained listeners was conducted. To avoid statistical bias in the results, each sentence pair was played twice in each test with the order of the sentences being reversed. Thus the total test comprized the comparison of some 800 sentence pairs. The results are shown in Tables 3 and 4.

Table 3
A/B comparison results for the FS1016 CELP coder

Speaker gender	SMWLPC [%]	STD LPC [%]
Female	59.5	40.5
Male	53.5	46.5
Total	56.5	43.5

An alternative view of the results is to look at the majority listener preference for the particular sentences. These results are shown in Table 5.
The results clearly indicate a preference for the SMWLPC coded speech in all instances and for both coders. This clear preference is despite the fact that the coding structures for both coders were left unaltered. Modifying the quantization procedures for the residual signal to suit the SMWLPC characteristics by, for example, retraining codebooks and introducing search weighting functions that suit the SMWLPC characteristics such as that proposed in [13] could be expected to show further substantial improvements in the performance of the coders when using SMWLPC.
It is interesting to note that for the CELP coder the preference for female speakers using SMWLPC was higher than for males and for the WI coder this was reversed. It is a well known property that CELP coders sound better for male speakers due to the retention of phase (temporal) information but poor modeling of the harmonic structure in the coding process [23]. Conversely harmonic type speech coders such as WI coders are better suited to female speakers due to the retention of the harmonic structure but loss of the phase information [23]. It appears that by removing more of the perceptually important information from the input speech before the residual is coded, SMWLPC is able to overcome some of the short comings of a particular low rate coding algorithm.
The results presented have extended and support those reported in [16] and [24] where a significant preference for SMWLPC coded sentences was reported. In [16] mean opinion score (MOS) testing was conducted using a 2 kbps WI coder and the results showed an improvement in MOS score from 3.31 to 3.45 when the standard LPC was replaced by SMWLPC.

Table 4
A/B comparison results for the WI coder

Speaker gender	SMWLPC [%]	STD LPC [%]
Female	54.5	45.5
Male	57.5	42.5
Total	56	44

Table 5
Majority preferred sentences

	SMWLPC [%]	STD LPC [%]	No preference [%]
CELP	70	30	0
WI	60	20	20
Total	65	25	10

5. Conclusion

A new technique which modifies the calculation of the LPC to better model the source for low rate speech coding has been developed. The technique involves the use of a psychoacoustic model to determine the simultaneously masked frequencies and also the frequencies whose power falls below the threshold of hearing. This information is then used to weight the power spectrum of the input speech, producing a modified power spectrum that contains only unmasked

information. A modified autocorrelation function is then generated via a DFT operation and standard recursive algorithms are used to solve for the LPC. Retaining the use of the standard recursive algorithms limits any increase in computational complexity and also ensures that a stable all pole filter is produced.
Experimental results have shown that the technique better models the spectrum in the unmasked formant regions and thus removes more of the perceptually important information from the input speech signal than a standard LP filter. Subjective listening tests using both CELP and WI coders has confirmed that this property improves the perceptual quality of the synthesized speech for a given residual coding method.

Acknowledgements

Jason Lukasiak is in receipt of an Australian Postgraduate Award (Industry) and a Motorola (Australia) partnerships in research grant. Whisper Laboratories is funded by Motorola and the Australian Research Council.

References

[1] B. C. J. Moore, *An Introduction to the Psychology of Hearing*. Sydney: Academic Press, 1997.

[2] H. W. Strube, "Linear prediction on a warped frequency scale", *J. Acoust. Soc. Am.*, vol. 68, no. 4, pp. 1071–1076, 1980.

[3] Y. Nakatoh, T. Norimatsu, A. Heng Low, and H. Matsumoto, "Low bit rate coding for speech and audio using mel linear predictive coding (MLPC) analysis", in *Proc. ICSLP*, 1998.

[4] H. Hermansky, "Perceptual linear predictive analysis of speech", *J. Acoust. Soc. Am.*, vol. 87, no. 4, pp. 1738–1753, 1990.

[5] "MPEG4", ISO/IEC FCD 14496-3.

[6] L. B. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*. New Jersey: Prentice Hall, 1978.

[7] A. M. Kondoz, *Digital Speech*. New York: Wiley, 1995.

[8] J. Makhoul and J. Wolf, "Linear prediction and the spectral analysis of speech", BBN report, no. 2304, August 1972.

[9] J. Makhoul, "Linear prediction: a tutorial review", *Proc. IEEE*, vol. 63, pp. 561–580, 1975.

[10] B. Scharf, "Critical bands", in *Foundations of Modern Auditory Theory*, J. Tobias, Ed. New York: Academic Press, 1970, pp. 159–202.

[11] M. Schroeder and B. S. Atal, "Predictive coding of speech signals and subjective error criteria", in *IEEE Trans. ASSP*, 1979, pp. 247–254.

[12] P. Kroon and E. F. Deprettere, "A class of analysis by synthesis predictive coders for high quality speech coding at rates between 4.8 and 16 kbits/s", *IEEE J. Selec. Areas Commun.*, vol. 62, pp. 353–363, 1988.

[13] D. Sen, D. H. Irving, and W. H. Holmes, "PERCELP – perceptually enhanced random codebook excited linear prediction", in *Proc. IEEE W/shop Speech Cod. Telecommun.*, 1993, pp. 101–102.

[14] I. S. Burnett, "Hybrid techniques for speech coding", Ph.D. thesis, University of Bath, 1992.

[15] J. D. Johnston, "Transform coding of audio signals using perceptual noise criteria", *IEEE J. Selec. Areas Commun.*, vol. 6, pp. 314–323, 1988.

[16] J. Lukasiak and I. S. Burnett, "Exploiting simultaneously masked linear prediction in a WI speech coder", in *Proc. IEEE W/shop Speech Cod.*, 2000, pp. 11–13.

[17] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing*. New Jersey: Prentice Hall, 1996.

[18] National Communication System, details to assist in implementation of Federal Standard 1016 CELP, Office of the manager National Communication System, Arlington.

[19] G. S. Kang and L. J. Fransen, "Low-bit rate speech encoders based on line spectrum frequencies (LSFs)", NRL report, no. 8857, Naval Research Lab., Washington D.C., Jan. 1985.

[20] M. Yong, G. Davidson, and A. Gersho, "Encoding of LPC spectral parameters using switched adaptive inter frame vector prediction", in *Proc. ICASSP*, 1988, vol. 1, pp. 402–405.

[21] K. K. Paliwal and B. S. Atal, "Efficient vector quantisation of LPC parameters at 24 bits/frame", *IEEE Trans. Speech Audio Proc.*, vol. 1, no. 1, pp. 3–14, 1993.

[22] W. B. Kleijn and J. Haagen, "A speech coder based on decomposition of characteristic waveforms", in *Proc. ICASSP*, 1995, vol. 1, pp. 508–511.

[23] J. Skoglund and W. B. Kleijn, "On time frequency masking in voiced speed", *IEEE Trans. Speech Audio Proc.*, vol. 8, no. 4, pp. 361–369, 2000.

[24] J. Lukasiak, I. S. Burnett, J. F. Chicharo, and M. M. Thomson, "Linear prediction incorporating simultaneous masking", in *Proc. ICASSP*, 2000, vol. 3, pp. 1471–1474.

Jason Lukasiak is currently a Ph.D. student in the Institute for Telecommunications Research at the University of Wollongong, Australia. He received a B.E. (Hons.) from the University of Wollongong in 1998 and immediately commenced studying for a Ph.D. He worked for BHP Slab and Plate products from 1987 to 1997 where his positions ranged from computer network technician to Electrical Project Engineer. During this time he studied part time receiving an electrical trades certificate, advanced certificate in computer technology and associate diploma of electrical engineering. Jason's Ph.D. research topic is scalable speech compression over a range of bit rates from $1-8$ kbps.
e-mail: j101@ouw.edu.au
Whisper Laboratories, TITR
University of Wollongong
Wollongong NSW 2522, Australia

Ian S. Burnett is Director of the Telecommunications Research Centre and a Senior Lecturer at the University of Wollongong, Australia. He received B.Sc. and M.Eng. degrees from the University of Bath, UK in 1987 and 1988, respectively. From 1987 he was with GEC Marconi Secure Radio working on digital communications and speech compression. In 1989 he returned to the University of Bath under a Vodafone scholarship, and completed a Ph.D. in "Hybrid Techniques for Speech Coding" in 1992. In 1993 he worked for Loughborough Sound Images on various real-time signal processing systems, before taking his present position. His current research interests lie primarily in speech/audio compression, audio scene analysis and multimedia. He is currently active in the ISO MPEG standardization concentrating on MPEG-21 and MPEG-7-Audio.
e-mail: j101@ouw.edu.au
Whisper Laboratories, TITR
University of Wollongong
Wollongong NSW 2522, Australia

Paper

Walsh-chirp sequences for wireless applications

Beata J. Wysocki, Tadeusz A. Wysocki, and Hans-Jürgen Zepernick

Abstract — This paper deals with a new method to design polyphase spreading sequences for DS CDMA wireless applications. The method is based on weighting symbols of the orthogonal Walsh sequences by the complex factors being symbols of baseband chirp sequences. The resulting sequences possess good aperiodic correlation properties, while maintaining the orthogonality. Because of the parametric design, the sequences can be optimized to achieve desired characteristics.

Keywords — spread spectrum communications, polyphase spreading sequences, multi-access interference, Walsh sequences, aperiodic correlation functions.

1. Introduction

Several families of complex spreading sequences have been proposed in literature, with some of them, e.g. [1] allowing for a good compromise between autocorrelation (AC) and cross correlation (CC) properties or even achieving orthogonality in the case of perfect synchronization [2]. Because for the downlink (base station to mobile transmission) the conditions for synchronous operation can be met, orthogonality would allow for cancellation of multi-access interference (MAI) for the down-link, and for simpler receivers in mobile terminals. This is not the case for an uplink transmission, where synchronization at the base station of arrival times of signals from different mobile terminals is very difficult, if possible to achieve.

An MAI impact on the system performance has been studied in [3], and theoretical formulae for an equivalent signal-to-noise ratio (SNR) approximation are given there. However, for short spreading sequences, these formulae can be regarded as very rough estimates, only. A reasonable approach to compare different sequence sets in such a case is to obtain error performance simulating the DS CDMA system under the same assumptions as those used in [3] for derivation of the SNR formulae.

In the paper, we propose a method to design sets of orthogonal polyphase sequences obtained by modification of orthogonal Walsh sequence. It is achieved by weighting symbols of the Walsh sequences by the complex factors obtained from the superposition of baseband chirp sequences [4]. Because of the parametric design, the Walsh-chirp sequences can be optimized to achieve the desired characteristics. In the numerical example, we show that the designed sequences can possess good aperiodic cross correlation and autocorrelation properties, enabling an increase in the number of simultaneous users compare to the number of users achievable, if Gold or Gold-like [5] sequences of the similar length are employed.

We also compare the performance of Walsh-chirp sequences with the sequences designed using the method described in [2]. For the same sequence length of 16, there are only 8 such sequences and up to 16 sequences using our method. In addition, the simulated bit-error-rate (BER) performance, once again indicate superiority of the Walsh-chirp sequences compared to the sequences designed using the method described in [2].

2. Poly-chirp sequences

For chirp modulation, an elementary phase pulse is given by [6]:

$$q_p(t) = \begin{cases} \frac{t^2}{2T^2} - \frac{t}{2T}, & 0 < t \leq T \\ 0, & \text{otherwise} \end{cases}, \tag{1}$$

where T is the duration of the pulse. Thus, a baseband chirp pulse $b(t)$ is of the form:

$$b(t) = \begin{cases} \exp\left[j2\pi h q_p(t)\right], & 0 < t \leq T \\ 0, & \text{otherwise} \end{cases}. \tag{2}$$

Discretizing the pulse $b(t)$ by substituting n for t, and N for T, we can write a formula defining a complex polyphase chirp sequence

$$\left\{\hat{b}_n(h)\right\} = \left(\hat{b}_n(h);\ \ n = 1, 2, \ldots, N\right), \tag{3}$$

where:

$$\hat{b}_n(h) = \exp\left[j2\pi h b_n\right];\ \ n = 1, 2, \ldots, N \tag{4}$$

$$b_n = \frac{n^2 - nN}{2N^2}, \tag{5}$$

with h being an arbitrary nonzero real constant.

The latest result can be generalized to obtain poly-chirp sequences. In order to do so, let us define a pulse referred to as a chirp pulse of the order s, if and only if the first time derivative of its instantaneous frequency (the angular acceleration) is a step function with the number of time intervals where it is constant being equal to s. Based on the above definition, we can derive the formulae for the elementary phase pulses for baseband chirps of any order. Again, substituting n for t, and N for T, we can get then formula defining a complex poly-chirp sequence of order s.

For example, a double-chirp sequence $\{d_n\}$ is given by [4]:

$$d_n = \begin{cases} \frac{2n^2}{N^2} - \frac{n}{N}, & 0 < n \le N \\ -\frac{2n^2}{N^2} + \frac{3n}{N} - 1, & \frac{N}{2} < n \le N \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

and the complex double chirp sequence elements $\hat{d}_n$ are therefore given by:

$$\hat{d}_n = \exp[j 2\pi h d_n]; \quad n = 1, 2, \ldots, N. \tag{7}$$

Another class of sequences can be obtained if a superposition of chirp sequences of different orders is used to create complex polyphase sequences.

3. Walsh-chirp sequences

Let us consider here the sequences, $\{\hat{o}_n^{(i)}\}$, $i = 1, 2, \ldots, M$ having their elements $\hat{o}_n^{(i)}$ given by:

$$\hat{o}_n^{(i)} = w_n \hat{h}_n^{(i)}, \quad n = 1, 2, \ldots, N, \tag{8}$$

where

$$w_n = \exp\left[j 2\pi (c_1 b_n + c_2 d_n)\right] , \tag{9}$$

b_n and d_n are defined by Eqs. (5) and (6), respectively, c_1, c_2 are two real constants, and $\hat{h}_n^{(i)}$ are the elements of orthogonal sequences $\{\hat{h}_n^{(i)}\}$, $i = 1, \ldots, M$.
Because of Eq. (9), we have

$$w_n w_n^* = 1; \quad n = 1, \ldots, N, \tag{10}$$

where w_n^* denotes the complex conjugate of w_n. Hence, it is easy to show that the sequences $\{\hat{o}_n^{(i)}\}$ are also orthogonal, as long as the factors w_n are kept constant for $i = 1, 2, \ldots, M$.
Therefore, choosing the sequences $\{\hat{h}_n^{(i)}\}$ as the sequences obtained from the orthogonal Walsh functions, we can produce a set of the orthogonal complex Walsh-chirp sequences. The cross- and autocorrelation performance of the set depends not only on the Walsh functions chosen in the first place, but also on the values of the parameters c_1 and c_2. In the next Section, we show that these parameters can be optimized to achieve the desired characteristics.

4. Numerical example

In order to design a set of Walsh-chirp sequences of length 16, let us consider a set of 13 Walsh sequences given in Table 1. Then, we find the values of the coefficients c_1 and c_2 which minimize the mean square value of the aperiodic CC (MSACC) [1] for the whole set.
For that purpose, we calculated the MSACC for $0 \le c_1 \le 30$ and $0 \le c_2 \le 30$, with the grid of 0.2. In the investigated region, it reaches the minimum of 0.8532 for $c_1 = 15.8$ and $c_2 = 24.4$. For those values of c_1 and c_2 the mean square value of the aperiodic AC (MSAAC) [1] is equal to 1.5962. The orthogonality of the designed sequence set is clearly visible in Fig. 1, where we plotted two example aperiodic CC functions (ACCFs).

Table 1
Set of 13 Walsh sequences

No.	Binary spreading sequence
1	- - - - + + + + - - - - + + + +
2	- - + + + + - - - - + + + + - -
3	+ + - - - - + + - - + + + + - -
4	+ + - - + + - - - - + + - - + +
5	- - + + - - + + - - + + - - + +
6	- + + - - + + - - + + - - + + -
7	+ - - + + - - + - + + - - + + -
8	+ - - + - + + - - + + - + - - +
9	- + + - + - - + - + + - + - - +
10	- + - + + - + - - + - + + - + -
11	+ - + - - + - + - + - + + - + -
12	+ - + - + - + - - + - + - + - +
13	- + - + - + - + - + - + - + - +

To assess the usefulness of the designed sequence set, we simulated operation of the DS CDMA system utilizing these sequences. We have assumed that data transmitted in any of the active channels is random, grouped into 1000 packets of 524 bits. The system has been considered as an asynchronous one, with only the examined channel kept synchronized to the corresponding reference sequence generated in the receiver, while the interfering m channels have been randomly delayed with respect to the examined channel.

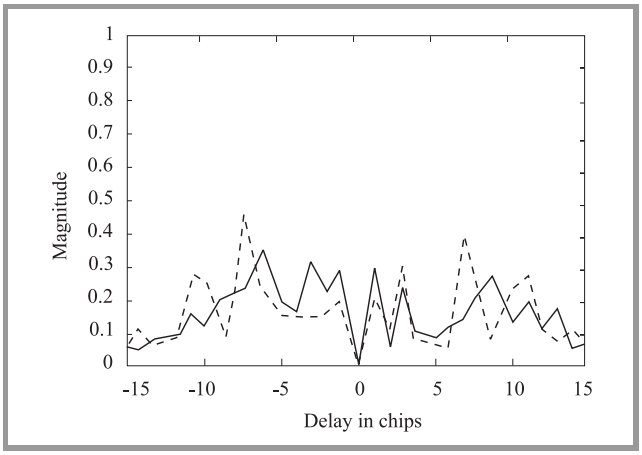


Fig. 1. Magnitudes of the ACCFs between the sequences: (1,10) – solid line, and (4,11) – dashed line.

Those random delays, τ_i, $i = 1, \ldots, m$, have been chosen as integer multiplies of 0.5, and satisfying the condition:

$$0 \le \tau_i < N. \tag{11}$$

Since in a real system phases of the generators used in all of the transmitting terminals can be different, we multiplied each of the interferers' signals by a coefficient:

$$\rho_i = \exp(j\phi_i), \tag{12}$$

where ϕ_i is a constant chosen randomly from the interval $[0, 2\pi]$.

In order to simplify the simulations, we have kept those randomly chosen coefficients, τ_i, ϕ_i, $i = 1, \ldots, m$, constant throughout the transmission of a single packet in the examined channel, with drawing of them repeated before simulation of every new transmission of a single packet.

For each of the simulated packet transmissions in the examined channel, the sequences used by the interferers has been chosen randomly from the set of all possible spreading sequences utilized in the system, disregarding the one used by the channel under examination.

Apart from the presence of MAI, we have assumed the presence of white Gaussian noise in the channel. We performed the simulations for $E_b/N_0 = 20$ dB, and $E_b/N_0 = 8$ dB, where N_0 is single sided power spectral density of white nose, and E_b is energy per information bit. To avoid being drawn into considering the problems associated with the "near-far-effect", we have assumed powers of all signals arriving at the receiver kept at the same level.

The achieved BER is plotted in Fig. 2 versus the number of interfering channels. For the comparison, we present there also BER characteristic obtained in the case when the designed sequence set is replaced by the set of 15-chip Gold-like sequences [5].

From the results presented in Fig. 2, it is clearly visible that the system utilizing Walsh-chirp sequences significantly outperforms the one utilizing Gold-like sequences, allowing for considerably more simultaneous users in the system at the same level of BER.

Then, we have compared performance of the Walsh-chirp sequence with the orthogonal polyphase spreading sequences proposed in [2] for the length $N = 16$. The set of sequences in [2] is defined by $U_{m,n}(N) = \{u_m : 1 \le M < N\}$, while the ith element of a given sequence u_M is defined by

$$u_M(i) = (-1)^{M_i} \exp\left[\frac{j\pi(iM^m + i^n)}{N}\right]; \quad 1 \le i < N, \tag{13}$$

where m is any positive nonzero integer, n is a real number, while M and N are relatively prime numbers.

In our case, $N = 16$, the parameter M can take only the values $M \in \{1, 3, 5, 7, 9, 11, 13, 15\}$, so the maximum number of sequences is equal to 8. There is no general method given in [2] for finding the values of the parameters m, and n. Following [2], we chose the value of $m = 1.0$, and calculated MSACC and the MSAAC as functions of the parameter n. The obtained plots are given in Fig. 3.

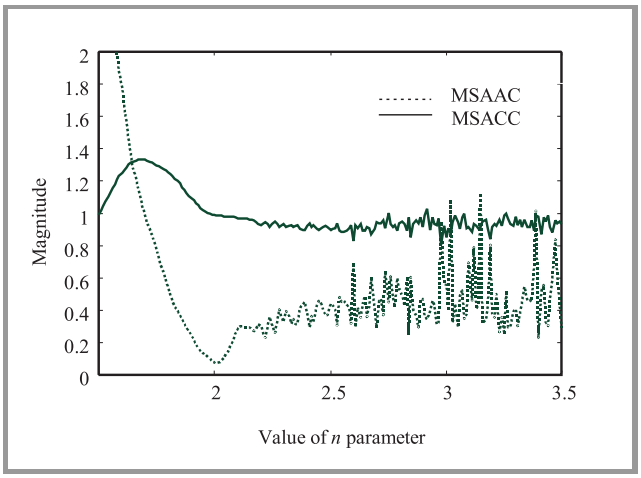


Fig. 3. Plots of MSAAC and MSACC as a function of the parameter n for the sequence set defined using formula (13).

For simulating BER performance we took $n = 2.5$ where there is a reasonable compromise between the values of MSACC and MSACC equal to 0.8968 and 0.4394, respectively. The achieved BER is plotted in Fig. 4 versus the number of interfering channels. Once again, it is visible

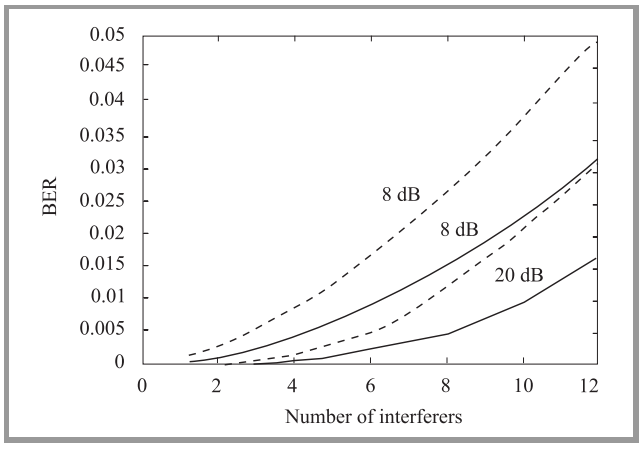


Fig. 2. BER as a function of the number of interfering channels; 16-chip Walsh-chirp sequences – solid lines, 15-chip Gold-like sequences – dashed lines.

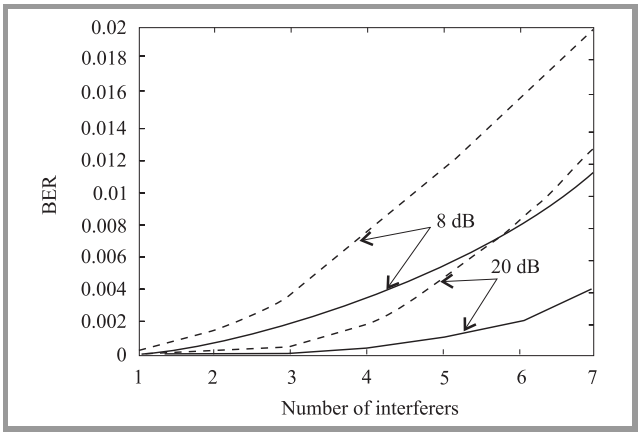


Fig. 4. BER as a function of the number of interfering channels; 16-chip Walsh-chirp sequences – solid lines, 16-chip sequences designed according to [2] – dashed lines.

that the Walsh-chirp sequences outperform sequences proposed in [2] for $N = 16$.

5. Conclusions

In the paper, we introduced a new method to design sets of orthogonal polyphase sequences. The method is based on utilizing a binary orthogonal sequence set and weighting the symbols of binary sequences with the complex coefficients obtained from the superposition of baseband chirp sequences. The resultant polyphase sequences can be optimized to achieve desired correlation properties of the set. As indicated by the simulations, error performance of a DS CDMA system utilizing the sequence set developed in the numerical example is significantly better than in the case of a system employing Gold-like sequences of a similar length, or orthogonal polyphase sequences of the same length, $N = 16$. Further on, one can apply the method proposed in this letter together with the sequences proposed in [2], i.e. instead of Walsh sequences, the sequences proposed in [2] can be modified using chirp pulses.

References

[1] I. Oppermann and B. S. Vucetic, "Complex spreading sequences with a wide range of correlation properties", *IEEE Trans. Commun.*, vol. COM-45, pp. 365–375, 1997.

[2] I. Oppermann, "Orthogonal complex-valued spreading sequences with a wide range of correlation properties", *IEEE Trans. Commun.*, vol. COM-45, pp. 1379–1380, 1997.

[3] M. N. Pursley, "Performance evaluation for phase-coded spread-spectrum multiple-access communication". Part I: "System analysis", *IEEE Trans. Commun.*, vol. COM-25, pp. 795–799, 1977.

[4] B. J. Wysocki and T. A. Wysocki, "Chirp sequences for wireless data networks", in *Proc. 5th Int. Symp. Commun. Theory Appl.* (ISCTA'99), Ambleside, U.K., 11–16 July 1999, pp. 202–204.

[5] P. Fan and M. Darnell, *Sequence Design for Communications Applications*. New York: Wiley, 1996.

[6] T. Wysocki, "Chirp modulation", in *Encyclopedia of Electrical and Electronics Engineering*, J. G. Webster, Ed. New York: Wiley, 1999, vol. 13.

Beata Joanna Wysocki graduated from Warsaw University of Technology receiving her M.Eng. degree in electrical engineering in 1991. In 1994, she started the Ph.D. study in the Australian Telecommunications Research Institute at Curtin University of Technology. In March 2000, she was awarded her Ph.D. for the thesis: "Signal Formats for Code Division Multiple Access Wireless Networks". During the Ph.D. studies, she was involved in a research project "Wireless ATM Hub" at the Cooperative Research Centre for Broadband Telecommunications and Networking, and worked as a research assistant at Edith Cowan University, within the ARC funded "CDMA with enhanced protection against frequency selective fading", and "Reliable high rate data transmission over microwave local area networks". Since October 1999 she has been with the Telecommunications & Information Technology Research Institute at the University Wollongong as a research fellow. Her research interests include sequence design for direct sequence (DS) code division multiple access (CDMA) data networks and error control strategies for broadband wireless access (BWA) systems.
e-mail: beata@snrc.uow.edu.au
University of Wollongong, T.I.T.R.
Northfields Av.
Wollongong NSW 2522, Australia

Tadeusz Antoni Wysocki received the M.Sc.Eng. degree with the highest distinction in telecommunications from the Academy of Technology and Agriculture, Bydgoszcz, Poland, in 1981. In 1984, he received his Ph.D. degree, and in 1990, was awarded a D.Sc. degree (habilitation) in telecommunications from the Warsaw University of Technology. In 1992, Dr. Wysocki moved to Perth, Western Australia to work at Edith Cowan University. He spent the whole 1993 at the University of Hagen, Germany, within the framework of Alexander von Humboldt Research Fellowship. After returning to Australia, he was appointed a Program Leader, Wireless Systems, within Cooperative Research Centre for Broadband Telecommunications and Networking. Since December 1998 he has been working as an Associate Professor at the University of Wollongong, NSW, within the School of Electrical, Computer and Telecommunications Engineering being a Program Leader of the Smart Networks Program within the Cooperative Research Centre for Smart Internet Technology. The main areas of Dr. Wysocki's research interests include: indoor propagation of microwaves, code division multiple access (CDMA), digital modulation and coding schemes, as well as mobile data protocols including those for ad hoc networks. He is the author or co-author of three books, over 100 research publications and nine patents. He is a Senior Member of IEEE.
e-mail: tad_wysocki@uow.edu.au
University of Wollongong
School of Electrical, Computer
and Telecommunications Engineering
Northfields Av.
Wollongong NSW 2522, Australia

Hans-Jürgen Zepernick received the Dipl.-Ing. degree form the University of Siegen, Germany, in 1987, and the Dr.-Ing. degree from the University of Hagen, Germany, in 1994, both in electrical engineering. From 1988 to 1989, he was a Research Engineer with the Radio and Radar Department at Siemens AG, Munich, Germany. Here he developed signal processing algorithms that are suitable for miniaturization of Bragg cell receivers. From 1989 to 1995, he was a Research Associate with the Institute of Telecommunications at the University of Hagen, where his work was focused on methods for

performance analysis of error control coding schemes with application to mobile communication systems. In May 1995, Dr. Zepernick joined the Cooperative Research Centre for Broadband Telecommunications and Networking (CRC-BTN) in Perth, Australia, as a Research Fellow and conducted research on physical layer and data link layer topics for wireless ATM. Since 1997, he has been a Project Leader for the Coding and Modulation project at the CRC-BTN. Since 1999, Dr. Zepernick has been employed as a Program Leader for the Wireless Program at the Australian Telecommunications Cooperative Research Centre (ATCRC), Perth, Australia. In August 2000, he has been appointed the Deputy Director of the Australian Telecommunications Research Institute (ATRI), Perth, Australia. Dr. Zepernick is an author or co-author of some 50 technical papers. His research interests include radio channel characterization and modeling, coding and modulation, equalization, spread-spectrum systems, wireless networks and future generation wireless systems.

e-mail: hans@atri.curtin.edu.au
Curtin University of Technology, ATRI
GPO Box U1987, Perth WA 6845, Australia

Paper

Memory truncation and crosstalk cancellation for efficient Viterbi detection in FDMA systems

Alfred Mertins

Abstract — In this paper, the design of optimal receive filter banks for frequency division multiple access (FDMA) over frequency selective channels is investigated. A new design strategy based on the principle of memory truncation, rather than equalization, is presented. Through the receive filters, each subchannel is truncated to a pre-defined length, and the final data recovery is carried out via low complexity Viterbi detectors. Both closed form designs and adaptive techniques are discussed. Design examples are presented for high speed transmission over copper wires. The examples show that memory truncation allows significant performance improvements over the often used minimum mean squared error (MMSE) equalization.

Keywords — transmultiplexer, memory truncation, Viterbi detector, dispersive channel, filter banks.

1. Introduction

The performance of transmission systems based on discrete multitone (DMT) modulation [1 – 3] or orthogonal frequency division multiplex (OFDM) [4] degrades rapidly when the length of the channel impulse response exceeds the length of the guard interval, which is introduced to cope with non-ideal channels. As a result of an insufficient guard interval, intersymbol interference (ISI) and inter-channel interference (ICI, crosstalk) will occur. To cope with longer channel impulse responses one can increase the length of the guard interval, but this will decrease the efficiency, as less data symbols can be transmitted. Increasing both the length of the guard interval and the number of subchannels allows one to maintain a desired bandwidth efficiency, but this strategy also has its limits. For example, the delay between transmitter and receiver may become unacceptably high. Also, the hardware requirements increase with an increasing number of subchannels. Finally, channels which can be regarded as slowly time-varying when the number of subchannels is low may turn into fast time-varying ones if the number of subchannels and thus the lengths of the transmit and receive filters are significantly increased.

In this paper, new methods for the design of optimal receive filter banks in multichannel transmission systems are proposed. The techniques are presented for a multirate filter bank framework, which gives a common description of a variety of transmission techniques [5]. The solutions apply to DMT [3, 6], OFDM [4], coded-OFDM [7], transmultiplexers [8, 9], and other transmission techniques where the transmit signal is created as a weighted linear combination of basis sequences with the data symbols being the weights. Even code division multiple access (CDMA) [10, 11] can be seen as a multirate filter bank. Figure 1 shows the general structure of the transmit/channel/receive model used in this paper. Depending on the actual modulation technique (DMT, OFDM, CDMA, etc.), the upsampling factor N, the number of subchannels, M, and the impulse responses $g_k(n)$ and $h_k(n)$ are chosen.

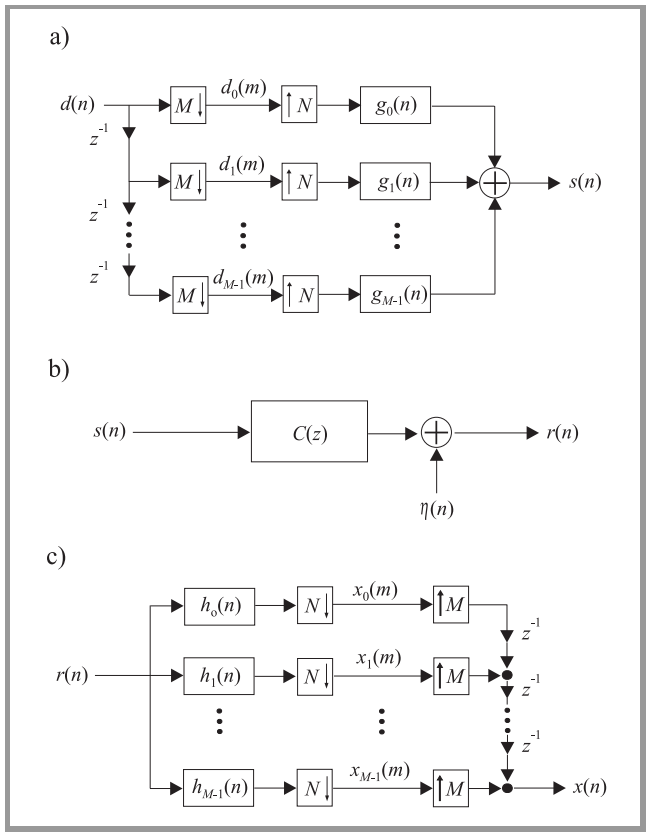


Fig. 1. Multirate discrete-time transmitter/channel/receiver model: (a) transmitter; (b) channel; (c) receiver.

Classically, the system in Fig. 1 is used to convert time-division multiplexed (TDM) data signals into frequency

division multiplexed (FDM) signals and vice versa. The synthesis filter bank in Fig. 1a then provides the TDM-to-FDM conversion. The FDM signal is transmitted through the channel, and the receive signal is finally fed into the analysis filter bank in Fig. 1c which converts the FDM signal back into a TDM one [8, 9]. Examples are DMT and OFDM, which both use block wise DFT's to do the TDM-to-FDM and FDM-to-TDM conversions. For DMT and OFDM the impulse responses of the filters in Fig. 1 are the complex exponentials occurring in the DFT, with maximum filter length N. A better frequency selectivity can be obtained if one uses longer filters, designed according to multirate filter bank theory. In [12] a comparison between cosine-modulated filter banks and DMT was given which shows that filter banks offer greater potential than block transforms.

Various solutions to the problem of reducing ISI through channel equalization have been proposed [5, 10, 12 – 19]. Most of them are based on minimizing the mean squared error (MSE) between the sent data and the equalizer output, either using a general MSE or a zero forcing (ZF) concept. Decision feedback equalizers (DFE) have been considered in [17]. Minimum mean squared error and ZF solutions with a joint design of receiver and transmitter have been proposed in [5, 16, 18]. Such a joint design can be useful in cases where communication takes place in both directions. In this paper, we concentrate on the optimal receiver design, thus addressing cases where the transmitter is fixed. The methods proposed in this paper are extensions of the technique in [19] to the design of entire receive filter banks for the critically sampled and oversampled cases (i.e $N \geq M$). Furthermore, methods for adaptive receiver design are presented. The design criterion is based on the idea of memory truncation [20, 21], where the receiver does not try to fully equalize the channel and leaves a residual system in the data path. In the optimum, the MSE between the equalizer output and a filtered version of the input data sequence is minimized. The final data detection then takes place via a Viterbi detector which needs to consider only the residual impulse responses. The lengths of the residual filters can be chosen arbitrarily and will typically be a few taps, thus allowing the use of low-complexity Viterbi detectors. The advantage of memory truncation over equalization is that critical channel zeros (e.g. zeros close to or even on the unit circle) need not be equalized, so that the problem of noise amplification through the equalizer can be avoided.

Note that for DMT transmission, memory truncation has also been proposed in a different form where the channel memory is shortened to the length of the guard interval prior to the DFT analysis in the receiver [22, 23]. In the present paper, however, memory truncation is incorporated as a property of the receive filters, and we can even treat cases where no guard interval is introduced at all.

The paper is organized as follows. In Section 2 the input-output relations for the multirate system in Fig. 1 are discussed. Section 3 addresses the design of optimal receive filter banks. Methods for adaptive receiver design are presented in Section 4. Results are discussed in Section 5, and finally, conclusions are given in Section 6.

Notation: The superscript ^{T} denotes transposition of a vector or matrix. The superscripts * and ^{H} denote complex conjugation and conjugate transposition $\left(\boldsymbol{r}^H = [\boldsymbol{r}^*]^T\right)$, respectively. $\boldsymbol{I}$ is an identity matrix of appropriate size. $E\{\ \}$ denotes the expectation operation, and $\delta_{i,k}$ is the Kronecker symbol.

2. Input-output relations

We consider the system in Fig. 1. The sequences $d_k(m)$, $k = 0, 1, \ldots, M-1$ are created through a series-to-parallel conversion of a single data sequence $d(m)$ in the form $d_k(m) = d(mM - k),\ k = 0, 1, \ldots, M-1$. In other words, they are polyphase components of the sequence $d(n)$. In the next step, the data sequences $d_k(m)$ are upsampled by a factor of N and then fed into the M respective synthesis filters with impulse responses $g_k(n),\ k = 0, 1, \ldots, M-1$. The sum of the filtered signals finally forms the transmit signal

$$s(n) = \sum_{k=0}^{M-1} \sum_{m=-\infty}^{\infty} d_k(m)\, g_k(n - mN)\ . \tag{1}$$

Typically, the filters $g_k(n)$ are chosen to be frequency selective, so that each data sequence $d_k(m)$ is transmitted in a distinct frequency band.

To make certain that the input data can be recovered at least theoretically from the transmit signal $s(n)$, the upsampling factor N must be chosen such that $N \geq M$ [9]. In many practical systems $N > M$ is used, which means that the transmitter introduces redundancy. This redundancy can be utilized in the receiver for enhancing the performance in the presence of frequency selective channels.

For the discussion in this paper, the transmission channel is assumed to be time-invariant. However, since adaptive methods for the receiver design are proposed the channel may, in practice, even be slowly time varying with respect to the filter lengths involved. Considering a time invariant channel, the receive signal is given by

$$r(n) = \left[\sum_{m=-\infty}^{\infty} c(m)\, s(n-m) \right] + \eta(n)\ , \tag{2}$$

where $\eta(n)$ is an additive, data independent noise process and $c(n)$ is the channel impulse response. The noise is assumed to be zero mean and wide-sense stationary.

On the receiver side, the signal $r(n)$ is fed into the analysis filter bank, as shown in Fig. 1, and the filter output signals are subsampled by a factor of N to form the final output signals

$$x_k(m) = \sum_{n=0}^{L_h-1} h_k(n)\, r(mN - n), \qquad k = 0, 1, \ldots, M-1\ . \tag{3}$$

In Eq. (3), L_h is the length of the receive filters. Combining Eqs. (1), (2) and (3) we get the input-output relation

$$x_k(m) = \sum_{n=0}^{L_h-1} \sum_{\mu=-\infty}^{\infty} \sum_{i=0}^{M-1} \sum_{\ell=-\infty}^{\infty} h_k(n)\, c(\mu)\, d_i(\ell) \times \\ \times\, g_i(mN-n-\mu-\ell N) + \sum_{n=0}^{L_h-1} h_k(n)\, \eta(mN-n), \quad (4) \\ k = 0,1,\ldots,M-1\,.$$

Under ideal conditions where the analysis and synthesis filters of the transmission system form a perfect reconstruction (PR) filter bank and where the channel is noise free and ideal (i.e. $\eta(n) = 0 \,\forall\, n$ and $c(n) = \delta_{n,0}$) the transmit/receive system allows us to recover the data $d_k(m)$ without error:

$$x_k(m) = d_k(m-m_0)\,. \quad (5)$$

The term m_0 is the overall delay of the system. The PR conditions for the filter bank itself are

$$\sum_{n=0}^{L_h-1} h_k(n)\, g_i(mN-n) = \delta_{i,k}\, \delta_{m,m_0} \quad (6)$$

with $i,k = 0,\ldots,M-1$. A practical problem is that even transmitter/receiver systems satisfying Eq. (6) will be unable to perfectly recover the data if a non-ideal channel is introduced. Thus, the channel should be taken into account when designing the receive filter bank. Methods for this will be discussed in the next section. Since the channel is usually not known a priori in practice, adaptation rules will be presented in Section 4.

3. Design of optimal receive filter banks

In this section, we derive methods for the design of optimal receive filter banks. For this we define an error signal as the difference between the receiver output signals $x_k(m)$ and filtered versions of the data sequences:

$$e_k(m) = \sum_{j=0}^{L_h-1} h_k(j)\, r(mN-j) + \\ - \sum_{i=0}^{L_p-1} p_k(i)\, d_k(m-m_0-i)\,, \quad (7) \\ k = 0,\ldots,M-1\,.$$

The optimality criteria for the design of the M receive filters are the MSE's given by

$$Q_k = E\{|e_k(m)|^2\}, \quad k = 0,\ldots,M-1\,, \quad (8)$$

which are to be minimized under the energy constraints

$$\sum_{n=0}^{L_p-1} |p_k(n)|^2 = 1, \quad k = 0,1,\ldots,M-1\,. \quad (9)$$

The constraints (9) are needed to avoid the trivial solution $h_k(n) = 0,\ p_k(n) = 0$.

Note that the error measure (7) is different from the MSE as defined for conventional MSE equalizers [10, 12 – 15]. The idea behind the proposed approach is to truncate the channel memory and not to delete it completely. The impulse responses $p_k(m)$ are to be understood as residual impulse responses of arbitrarily chosen length L_p. Both the optimal residual systems $p_k(m)$ and receive filters $h_k(n)$ need to be found through minimization of Eq. (8).

Because of the existence of residual systems $p_k(m)$, minimizing Eq. (8) does, in general, not result in an equalization of the channel. Even if $Q_k = 0$ there will be a remaining ISI between L_p consecutive data samples in each of the subchannels. The crosstalk between different channels $i \neq k$ will be reduced as much as possible with FIR filters of the given length L_h.

With analysis filters designed through the minimization of Eq. (8) the overall system can be modeled with little error as a set of M independent channels with

$$x_k(m) = \left[\sum_{i=0}^{L_p-1} p_k(i)\, d_k(m-m_0-i)\right] + \eta_k'(m)\,, \quad (10) \\ k = 0,\ldots,M-1\,.$$

The modified noise processes $\eta_k'(m)$ contains the filtered and subsampled original noise and all modeling errors made by simplifying the real system to the form (10).

To recover the data, the signals $x_k(m),\ k = 0,\ldots,M-1$ are fed into M independently operating Viterbi detectors which have to consider the respective channels $p_k(m)$, $k = 0,\ldots,M-1$. Since the lengths of these channels are chosen arbitrarily, one can choose lengths which result in a manageable computational cost for the Viterbi detectors while maintaining a low noise variance at the detector inputs. Clearly, the longer the systems $p_k(m)$ are, the smaller the modeling errors in (10) and thus the smaller the variances $E\{|\eta_k'(m)|^2\}$ are. For $L_p = 1$ the Viterbi detectors degenerate to simple threshold detectors, at the expense of an increased noise variance compared to cases where $L_p > 1$.

Note that in the special case of $L_p = 1$, Eq. (8) states a standard MSE criterion, and the optimized analysis filters $h_k(n)$ can be regarded as MMSE equalizers. Then the proposed solution becomes equivalent to other known MMSE solutions [10, 12 – 15].

To obtain a compact formulation of the objective function, we now introduce the following vectors:

$$\boldsymbol{h}_k = \left[h_k(0),\ldots,h_k(L_h-1)\right]^T\,, \quad (11)$$

$$\widetilde{\boldsymbol{r}}(m) = \left[r(mN),\ldots,r(mN-L_h+1)\right]^T\,, \quad (12)$$

$$\boldsymbol{p}_k = \left[p_k(0),\ldots,p_k(L_p-1)\right]^T\,, \quad (13)$$

$$\boldsymbol{d}_k(m) = \left[d_k(m),\ldots,d_k(m-L_p+1)\right]^T\,. \quad (14)$$

We get

$$e_k(m) = \tilde{\boldsymbol{r}}^T(m)\boldsymbol{h}_k - \boldsymbol{d}_k^T(m)\boldsymbol{p}_k \,. \tag{15}$$

Using this notation the cost functions (8) can finally be written as

$$Q_k = \boldsymbol{h}_k^H \boldsymbol{R}_{rr}\boldsymbol{h}_k - \boldsymbol{h}_k^H \boldsymbol{R}_{rd}^{(k)}\boldsymbol{p}_k - \boldsymbol{p}_k^H \boldsymbol{R}_{dr}^{(k)}\boldsymbol{h}_k + \boldsymbol{p}_k^H \boldsymbol{R}_{dd}^{(k)}\boldsymbol{p}_k \tag{16}$$

with

$$\boldsymbol{R}_{rr} = E\left\{\boldsymbol{r}^*(m)\ \boldsymbol{r}^T(m)\right\} \,,$$

$$\boldsymbol{R}_{rd}^{(k)} = \left[\boldsymbol{R}_{dr}^{(k)}\right]^H = E\left\{\boldsymbol{r}^*(m)\,\boldsymbol{d}_k^T(m-m_0)\right\} \,,$$

$$\boldsymbol{R}_{dd}^{(k)} = E\left\{\boldsymbol{d}_k^*(m-m_0)\,\boldsymbol{d}_k^T(m-m_0)\right\} \,.$$

For the sake of simplicity, let us assume that all data sequences $d_k(m)$ are spectrally white and have the same variance σ_d^2. Then the autocorrelation matrices $\boldsymbol{R}_{dd}^{(k)}$, $k = 0, \ldots, M-1$ are diagonal with diagonal entries σ_d^2,

$$\boldsymbol{R}_{dd}^{(k)} = \sigma_d^2 \boldsymbol{I} \,, \tag{17}$$

and Eq. (16) simplifies to

$$Q_k = \boldsymbol{h}_k^H \boldsymbol{R}_{rr}\boldsymbol{h}_k - \boldsymbol{h}_k^H \boldsymbol{R}_{rd}^{(k)}\boldsymbol{p}_k - \boldsymbol{p}_k^H \boldsymbol{R}_{dr}^{(k)}\boldsymbol{h}_k + \sigma_d^2 \boldsymbol{p}_k^H \boldsymbol{p}_k \,. \tag{18}$$

We now consider the minimization of Eq. (18) with respect to $\boldsymbol{p}_k$ and $\boldsymbol{h}_k$ under the energy constraints (9). To derive the optimal filters we first derive the optimal vector $\boldsymbol{h}_k$ given a fixed residual system $\boldsymbol{p}_k$. From $\partial Q_k / \partial \boldsymbol{h}_k = 0$ with Q_k as in Eq. (18), we get

$$\boldsymbol{h}_k^{(\mathrm{opt})} = \boldsymbol{R}_{rr}^{-1} \boldsymbol{R}_{rd}^{(k)}\, \boldsymbol{p}_k \,. \tag{19}$$

Substituting $\boldsymbol{h}_k^{(\mathrm{opt})}$ into Eq. (18) results in

$$Q_k = -\boldsymbol{p}_k^H \left[\boldsymbol{R}_{rd}^{(k)}\right]^H \boldsymbol{R}_{rr}^{-1} \boldsymbol{R}_{rd}^{(k)}\, \boldsymbol{p}_k + \sigma_d^2\, \boldsymbol{p}_k^H\, \boldsymbol{p}_k \,, \tag{20}$$

which now is to be minimized with respect to $\boldsymbol{p}_k$ under the constraint (9). This yields the eigenvalue problems

$$\left[\sigma_d^2 \boldsymbol{I} - \left[\boldsymbol{R}_{rd}^{(k)}\right]^H \boldsymbol{R}_{rr}^{-1} \boldsymbol{R}_{rd}^{(k)}\right] \boldsymbol{p}_k = \lambda_k \boldsymbol{p}_k, \qquad k = 0, 1, \ldots, M-1 \tag{21}$$

which are essentially similar to the one in [20] for the single-channel case. The optimal vectors $\boldsymbol{p}_k$ are the eigenvectors that belong to the respective smallest eigenvalues λ_k, $k = 0, \ldots, M-1$.

The receive filters designed according to the method described above minimize the error measures Q_k under the energy constraint and thus maximize the signal-to-noise ratios (SNR's) at the filter outputs. Since the filter output signals, together with the residual systems, are fed into the Viterbi detectors, the algorithm maximizes the SNR's as seen by the Viterbi detectors.

4. Adaptive receiver design

The receiver design method presented in the previous section may be difficult to implement under real-world conditions where the required computational power and accuracy are not available. Also, a real-world channel may be slowly time varying, which causes problems for the receiver design above. To avoid such problems we now derive adaptation rules for the receiver design. For this we follow the strategy for the single-channel case in [20]. During adaptation, we assume that the data sequences $d_k(m)$ are known or have been correctly estimated by the receiver, and we use the received samples as noisy estimates of the required correlation terms. We first look at the design of filter $h_k(n)$ (vector $\boldsymbol{h}_k$) according to the rule

$$\boldsymbol{h}_k^{(\mu+1)} = \boldsymbol{h}_k^{(\mu)} - \gamma_h\, e_k(\mu)\, \tilde{\boldsymbol{r}}_k^*(\mu) \,, \tag{22}$$

where μ denotes the iteration step. The value $e_k(\mu)$ is the value of the error defined in Eq. (7) in the μth iteration step, and $\tilde{\boldsymbol{r}}_k(\mu)$ is the receive vector in the μth step. Finally, γ_h is a factor that controls the step size and convergence speed. Thus, Eq. (22) is similar to the well-known LMS adaptation rule for equalizer design.

A rule for adapting $\boldsymbol{p}_k$ can be stated as

$$\boldsymbol{q}_k^{(\mu+1)} = \boldsymbol{p}_k^{(\mu)} + \gamma_p\, e_k(\mu)\left[\boldsymbol{d}_k(\mu)\right]^* \,, \tag{23}$$

$$\boldsymbol{p}_k^{(\mu+1)} = \frac{\boldsymbol{q}_k^{(\mu+1)}}{\left\|\boldsymbol{q}_k^{(\mu+1)}\right\|} \,. \tag{24}$$

The entire iteration is given by Eqs. (22), (23) and (24), where the normalization step in Eq. (24) is needed to ensure that the energy condition (9) will be satisfied by the final filter $\boldsymbol{p}_k^{(\infty)}$. Using the same arguments as in [20] one can show that the iteration indeed converges to the MMSE solution where the final vector $\boldsymbol{p}_k^{(\infty)}$ is the eigenvector of $\left[\sigma_d^2 \boldsymbol{I} - \left[\boldsymbol{R}_{rd}^{(k)}\right]^H \boldsymbol{R}_{rr}^{-1} \boldsymbol{R}_{rd}^{(k)}\right]$ that corresponds to the smallest eigenvalue.

5. Results

To demonstrate the performance of the proposed algorithms, we consider data transmission over telephone lines in an ADSL/VDSL related setting [24]. Figure 2 shows the channel impulse response considered in this example. It is assumed that the channel noise is comprised of near and far end crosstalk as well as white Gaussian noise, resulting in the total power spectral density depicted in Fig. 3. We consider the use of a cosine-modulated filter bank for creating the transmit signal, which is an interesting alternative to blockwise DFT's as in DMT. In [12, 25] it was shown that such filter bank based systems offer greater potential than blockwise DFT's because of their longer impulse responses and better frequency selectivity. However,

they need equalization on the receiver side. In the present example, the transmit signal is synthesized via a 16-band cosine-modulated filter bank with extended lapped transform (ELT) prototype [26]. As in [12, 25] pulse amplitude modulation is used to create a real-valued transmit signal.

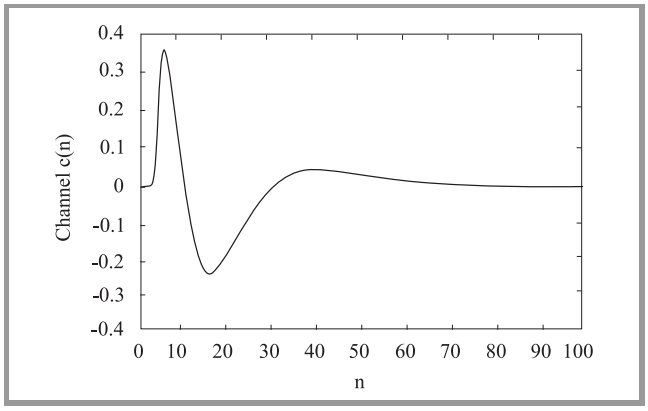


Fig. 2. Transmission channel impulse response.

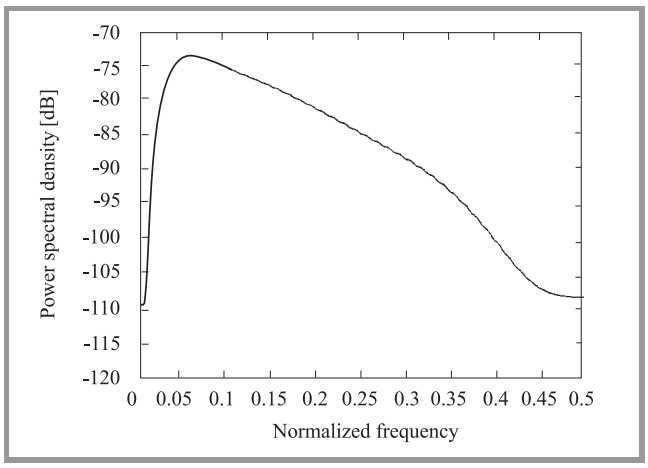


Fig. 3. Noise power spectral density.

Figure 4 shows the signal-to-noise ratios within the different bands at the equalizer output for several configurations. In all cases the lengths of the receive filters are chosen as $L_h = 128$. We first look at the results depicted in Fig. 4a. In this case, all bands are loaded with the same input power σ_d^2. This means that the transmission system is critically sampled and that no redundancy (e.g. in form of a guard interval) is introduced. The comparison of the three curves in Fig. 4a shows that, especially for the low-frequency channels, memory truncation ($L_p > 1$) significantly outperforms MMSE equalization ($L_p = 1$).
A significantly better performance of all methods under consideration is obtained if the first frequency band remains unloaded. Results are depicted in Fig. 4b. This strategy has been proposed in [25] as a possibility to introduce redundancy. Leaving out a particular band has two effects. Firstly, the system becomes oversampled, which means that the transmitter introduces redundancy in form

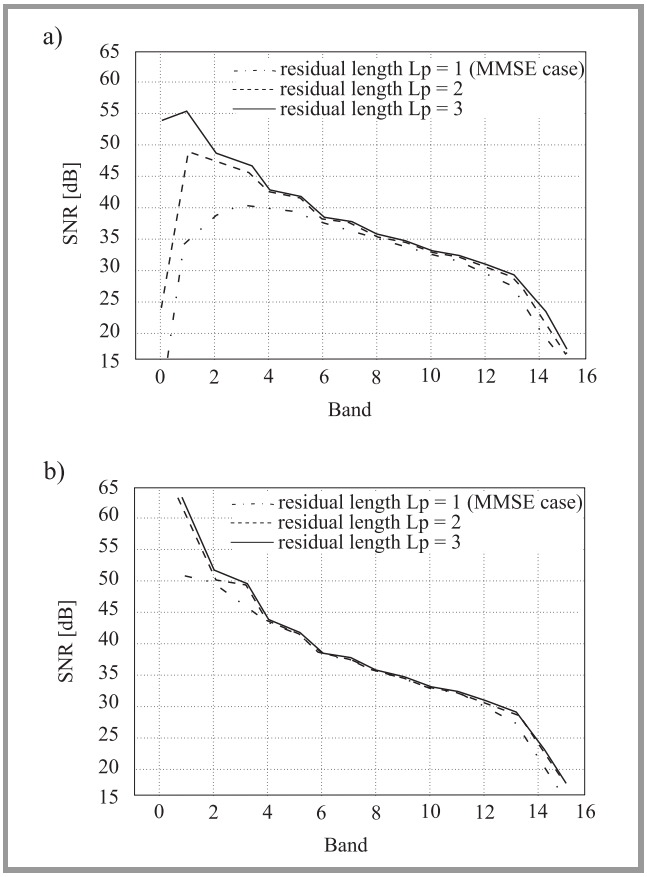


Fig. 4. Signal-to-noise ratios at detector input using a 16-band cosine-modulated filter bank as transmit filters: (a) all bands are loaded; (b) only bands $1-15$ are loaded.

of excess bandwidth. Secondly, the receive filters do not need to suppress crosstalk from the dropped channel and have more freedom to equalize their own data paths. As the results in Fig. 4b show, almost all channels gain from the fact that the first band has been left out. When comparing the three curves in Fig. 4b, we see that memory truncation still results in a noticeable improvement over MMSE equalization for a number of bands. The performance difference between $L_h = 2$ and $L_h = 3$, however, is only marginal in Fig. 4b.

From the above examples we see that a receiver based on memory truncation receive filters and low-cost Viterbi detectors can yield a significant improvement over MMSE equalization and threshold detection. In general, the amount of SNR improvement of memory truncation over MMSE equalization depends on the channel in question. Significant improvements can be expected whenever it is difficult to equalize a channel because of extreme frequency selectivity.

To demonstrate the adaptive approach, we consider the receive filter design for the first channel in the above setting. The residual filter length is chosen as $L_p = 3$. At the beginning of the iteration, the receive filter $h_1(n)$ was set to zero, and the residual system $p_1(n)$ was set

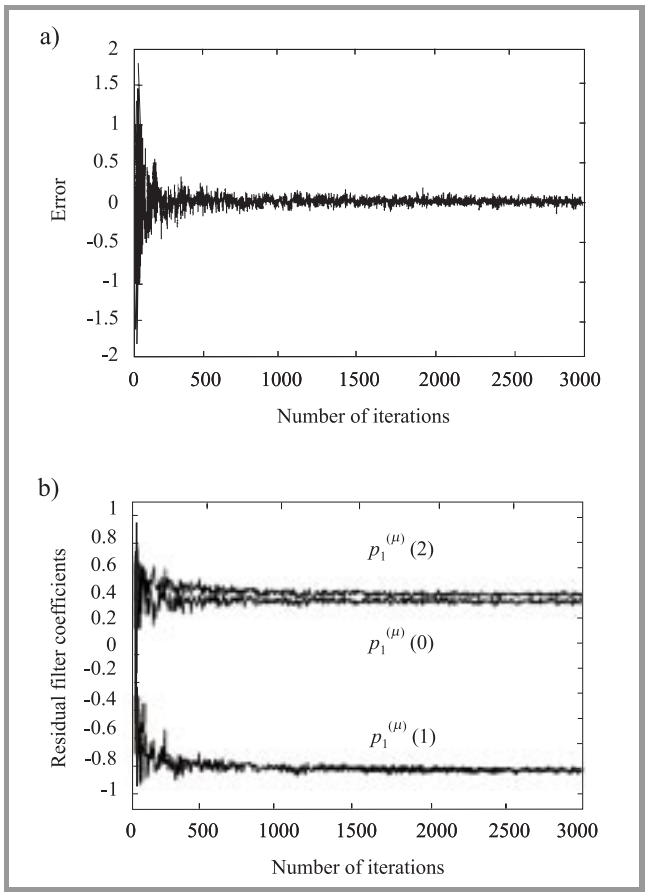


Fig. 5. Receiver adaptation: (a) error signal $e_1(\mu)$; (b) residual filter coefficients $p_1^{(\mu)}(m)$.

to $\{p_1^{(0)}(m)\} = \{1,0,0\}$. The step sizes were chosen as $\gamma_h = 0.0015$ and $\gamma_p = 0.33$. Figure 5a depicts the error signal $e_1(\mu)$, and it can be seen that the error rapidly decreases during the first few hundred iterations. Figure 5b shows the adaptation of the three residual coefficients. For comparison, the closed-form solution yields $\{p_1^{(\text{opt})}(m)\} = \{0.382, -0.822, 0.421\}$. These values are approached after some hundred iterations.

6. Conclusions

In this paper, optimal receive filter banks for FDM transmission systems have been presented. The receive filters are designed in such a way that the overall subchannel impulse responses become truncated to predefined lengths. Using an example of high-speed transmission over copper wires it was shown that the SNR can be significantly improved over MMSE equalizer banks. In general, the amount of improvement clearly depends on the channel in question, and there may be cases where MMSE approaches work equally well. The design methods presented are applicable to all transmultiplexing systems where the transmit signal is formed as a linear combination of transmit filter impulse responses with the data symbols being the weights (e.g. DMT, OFDM, CDMA). Extensions of the proposed methods to a joint transmitter/receiver design are under investigation.

References

[1] B. Hirosaki, "An analysis of automatic equalizers for orthogonally multiplexed QAM systems," *IEEE Trans. Commun.*, vol. 28, no. 1, pp. 73–83, 1980.

[2] L. J. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiple access," *IEEE Trans. Commun.*, vol. 33, no. 7, pp. 665–675, 1985.

[3] I. Kalet, "The multitone channel," *IEEE Trans. Commun.*, vol. 37, pp. 119–124, 1989.

[4] T. de Couasnon, R. Monnier, and J. B. Rault, "OFDM for digital TV broadcasting," *EURASIP Signal Proc.*, vol. 39, no. 1–2, pp. 1–32, 1994.

[5] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers". Part I: "Unification and optimal designs," *IEEE Trans. Signal Proc.*, vol. 47, no. 7, pp. 1988–2006, 1999.

[6] A. Ruiz, J. M. Cioffi, and S. Kasturia, "Discrete multiple tone modulation with coset coding for the spectrally shaped channel," *IEEE Trans. Commun.*, vol. 40, pp. 1012–1029, 1992.

[7] W. Y. Zou and Y. Wu, "COFDM: an overview," *IEEE Trans. Broadcast.*, vol. 41, pp. 1–8, 1995.

[8] M. G. Bellanger and J. L. Daguet, "TDM–FDM transmultiplexer: digital polyphase and FFT," *IEEE Trans. Commun.*, vol. 22, no. 9, pp. 1199–1205, 1974.

[9] M. Vetterli, "Perfect transmultiplexers," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Proc.*, Tokyo, Apr. 1986, pp. 48.9.1–48.9.4.

[10] M. K. Tsatsanis and G. B. Giannakis, "Optimal linear receivers for DS-CDMA systems: a signal processing approach," *IEEE Trans. Signal Proc.*, vol. 44, pp. 3044–3055, 1996.

[11] A. N. Akansu, P. Duhamel, X. Lin, and M. de Courville, "Orthogonal transmultiplexers in communication: a review," *IEEE Trans. Signal Proc.*, vol. 46, no. 4, pp. 979–995, 1998.

[12] L. Vandendorpe, L. Cuvelier, F. Deryck, J. Louveaux, and O. van de Wiel, "Fractionally spaced linear and decision-feedback detectors for transmultiplexers," *IEEE Trans. Signal Proc.*, vol. 46, no. 4, pp. 996–1011, 1998.

[13] R. P. Ramachandran and P. Kabal, "Transmultiplexers: perfect reconstruction and compensation of channel distortion," *Signal Proc.*, vol. 21, no. 3, pp. 261–274, 1990.

[14] L. Vandendorpe, "MMSE equalizers for multitone systems without guard time," in *Proc. 8th Eur. Signal Proc. Conf.*, Sept. 1996, vol. 3, pp. 2049–2052.

[15] T. Wiegand and N. J. Fliege, "Equalizers for transmultiplexers in orthogonal multiple carrier data transmission," *Ann. Telecommun.*, vol. 52, no. 1–2, pp. 39–45, 1997.

[16] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input–multiple-output (MIMO) transmission systems," *IEEE Trans. Signal Proc.*, vol. 42, no. 12, pp. 3221–3231, 1994.

[17] N. Al-Dhahir and J. M. Cioffi, "Block transmission over dispersive channels: transmit filter optimization and realization, and MMSE-DFE receiver performance," *IEEE Trans. Inform. Theory*, vol. 42, no. 1, pp. 137–160, 1996.

[18] T. Li and Z. Ding, "Joint transmitter-receiver optimization for partial response channels based on nonmaximally decimated filterbank precoding technique," *IEEE Trans. Signal Proc.*, vol. 47, no. 9, pp. 2407–2414, 1999.

[19] A. Mertins, "Memory truncation and crosstalk cancellation in transmultiplexers," *IEEE Commun. Lett.*, vol. 3, no. 6, pp. 180–182, 1999.

[20] D. D. Falconer and F. R. Magee, "Adaptive channel memory truncation for maximum likelihood sequence estimation," *Bell Syst. Tech. J.*, vol. 52, no. 9, pp. 1541–1562, 1973.

[21] K. D. Kammeyer, "Time truncation of channel impulse responses by linear filtering: a method to reduce the complexity of Viterbi equalization," *Archiv Elektrotech. Übertragungstech.*, vol. 48, no. 5, pp. 237–243, 1994.

[22] N. Al-Dhahir and J. M. Cioffi, "Optimum finite-length equalization for multicarrier transceivers," *IEEE Trans. Commun.*, vol. 44, no. 1, pp. 56–64, 1996.

[23] P. J. W. Melsa, R. C. Younce, and C. E. Rohrs, "Impulse response shortening for discrete multitone transceivers," *IEEE Trans. Commun.*, vol. 44, no. 12, pp. 1662–1672, 1996.

[24] K. Maxwell, "Asymmetric digital subscriber line: interim technology for the next forty years," *IEEE Commun. Mag.*, vol. 34, no. 10, pp. 100–106, 1996.

[25] S. Trautmann, T. Karp, and N. J. Fliege, "Using modulated filter banks for ISI/ICI-corrupted multicarrier transmission," in *Proc. SPIE Int. Symp. Opt. Sci., Eng., Instrum. Session: Wavelet Applications in Signal and Image Processing*, Denver, USA, July 1999.

[26] H. S. Malvar, "Extended lapped transforms: fast algorithms and applications," *IEEE Trans. Signal Proc.*, vol. 40, no. 11, pp. 2703–2714, 1992.

Alfred Mertins received his Dipl.-Ing. degree from the University of Paderborn, Germany, in 1984, the Dr.-Ing. degree in electrical engineering and the Dr.-Ing. habil. degree in telecommunications from the Hamburg University of Technology, Germany, in 1991 and 1994, respectively. From 1986 to 1991 he was with the Hamburg University of Technology, Germany, from 1991 to 1995 with the Microelectronics Applications Center Hamburg, Germany, from 1996 to 1997 with the University of Kiel, Germany, and from 1997 to 1998 with the University of Western Australia. Since September 1998 he has been with the University of Wollongong as a Senior Lecturer. His research interests include image and video processing, wavelets and filter banks, and digital communications.

e-mail: mertins@uow.edu.au
University of Wollongong
School of Electrical, Computer,
and Telecommunications Engineering
Northfields Avenue
Wollongong, NSW 2522, Australia

Paper

From modelling of a CDMA transceiver in indoor environment to an ASIC circuit synthesis

Sylvain Bourdel, Eric Campo, Patrick Melet, and Laurent Andrieux

Abstract — This paper presents the study, design and simulation of a multi-flow radio frequency transceiver based on a direct sequence spread-spectrum with a 2.4 GHz carrier. First, the functional model of differential QPSK modulation for digital transmission, and the different parts making up spread spectrum function (spreader, despreader, tracking and synchronising devices) have been studied, implemented, simulated and validated in noisy multi-users and multi-path environment by using a unified language. The results obtained by taking into account the home automation running constraints have allowed to determine some critical parameter values and so to integrate the digital functions in an ASIC circuit.

Keywords — indoor radiocommunications, wireless transceiver modelling, CDMA, DS-SS systems.

1. Introduction

The wireless communication needs are increasing for a few years in many application domains as home automation, automobile, aeronautic or telephony. The full expansion of telephony, for example, leads to the emergence of new communication techniques due to the need of higher data rate link. Thus, in 10 years we've passed from GSM-1 G (9.6 kbits) to UMTS-3 G (2 Mbits) [1] entailing a higher system complexity, and so a higher conception difficulty. From that, the telecommunication system conception is the more difficult since many different techniques (RF, numeric, analogic) must coexist in the same chip. Moreover, low cost integration and reduction of the conception time are a priority.

A good way to proceed is to start the conception by studying and analysing the system in its integrity and environment. Such approach allows the designer to choose the system architecture and optimise the parameters with the knowledge of the sub-system interactions (between them or with their environment). If this analysis is done with functional modelling and simulations under an appropriated software environment, it will entail a feedback between the integration stage and the conception as shown in Fig. 1. This top-down approach is retained for the realisation of our system. From the specifications a virtual prototype which allowed us to integrate the system on chip (SoC) has been realised.

The aim of this paper is to expose our solution for the conception of complex system based on a global approach. Here after, a system including its environment is described in a unified language. We will show how the system architecture can be chosen, how some system constraints can be highlighted thanks to such model and how the key parameters can be determined. This analysis allows us to define the optimum architecture for our system considering its physical and technological constraints. Once defined, it can be integrated on an ASIC circuit, performing thus the global conception of the SoC.

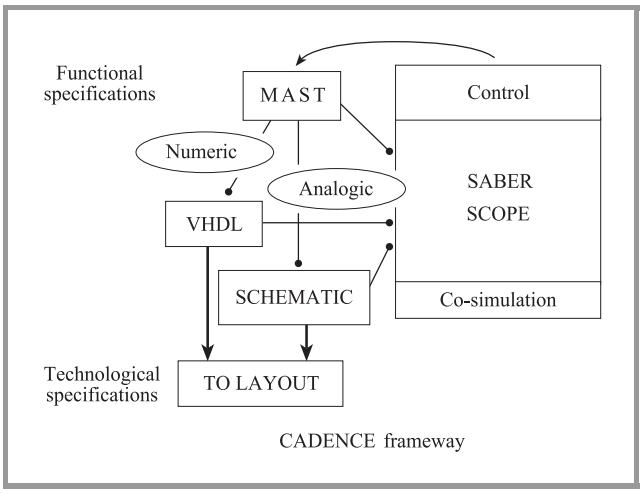


Fig. 1. Top-down conception platform using a unified language.

Based on this approach, the complete design of a 2.4 GHz CDMA transceiver for indoor communications will be presented from the functional simulation (with MAST language under SABER software) to the integration of the baseband processing functions on an ASIC with CADENCE tools. The direct sequence spread spectrum (DS-SS) technique is chosen because of its high indoor constraints immunity (fading, jamming...) [2]. Thanks to our global approach, the conception time is about 6 men/year for the design of the first ASIC.

In the first part, we present the global model of the transceiver. In the second part, we will present some simulation results, which allow us to validate our model, to choose architecture considering the system performances, to optimise some parameters and to highlight some system constraints. At last, a third part will present the integration of the baseband digital processing functions on an ASIC circuit. The conclusion will summarise the important results and the methodology used.

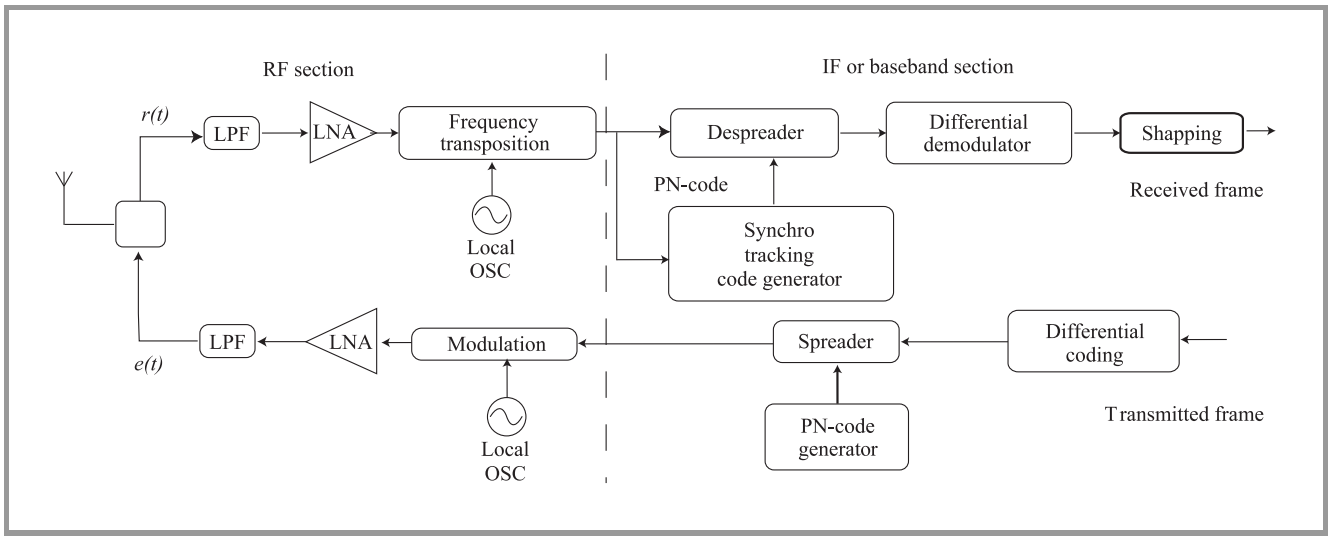


Fig. 2. Global principle scheme.

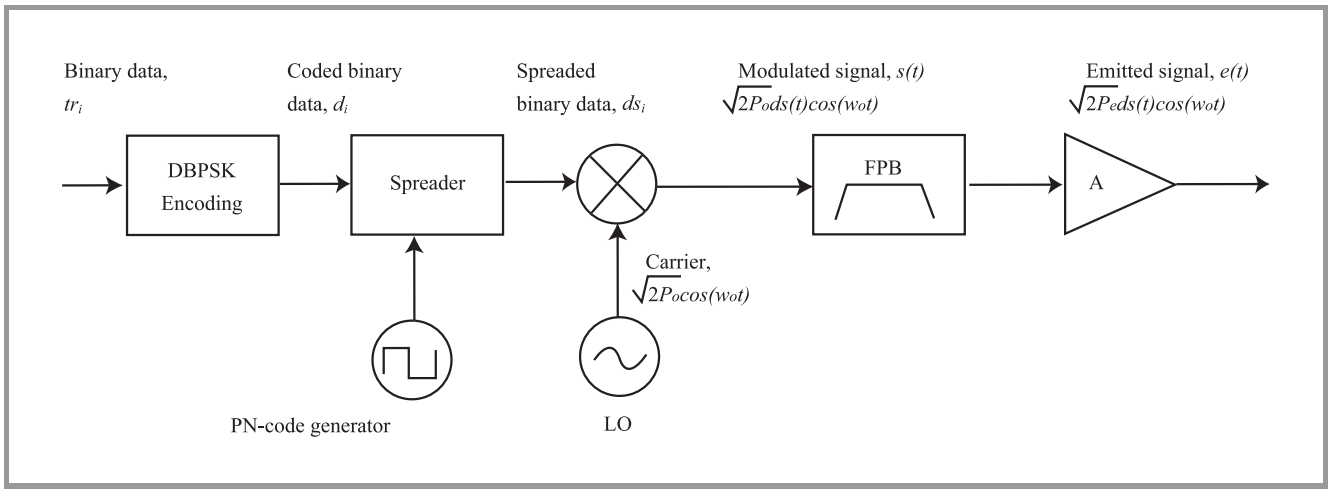


Fig. 3. DBPSK modulator model performed under SABER.

2. Modelisation of the complete transceiver with SABER

2.1. Principle

The system principle is shown in Fig. 2. All functions are described in MAST language which allows to simulate the RF stage and the transmission channels (AWGN, Rayleigh, Multiusers) in a simple way.

Several architectures have been studied and modelled, specially a baseband (BB) and an intermediate frequency (IF) one's. Also different synchronisation architectures (RASE and serial search) have been modelled in order to choose the most appropriated. The signal processing functions of the system can be divided into two groups: direct link and synchronisation. The following notation are used to describe the models: $d(t)$ – the transmitted information; d_i – the expression of the sampled sequence transmitted; $c(t)$ – the PN-code; $ds(t)$ – spread binary data; w_o – the carrier radian frequency; P_r – the power of the received signal; P_e – the power of the emitted signal; P_o – the power of the local oscillator (LO); Td – a time varying delay caused by the channel; T – the symbol time and Tc – the chip time.

2.2. Direct link elements

The modulation technique chosen is the differential phase shift keying (DPSK) for the following reasons [3]:

– an easiest realisation (because this modulation is non-coherent, that is to say there is no carrier tracking);

– the TEB $= 10^{-6}$ when $E_b/N_o = 11$ dB (at the demodulator input); the coherent modulation reaches the same TEB with $E_b/N_o = 12$ dB (see Fig. 8); this performance amelioration is not very significant in regard of the higher system complexity;

– the spectral efficiency can be increased ($\times 2$) just by using a QPSK modulation, which is easy to realise starting from a BPSK modulation [4].

The modulator model realised is presented in Fig. 3. The differential coding equation is [5]: $d_i = tr_i \oplus d_{i-1}$. In DS-SS, the correlation between the transmitted sequence $d(t)$ and the pseudo noise sequence $c(t)$, involving the signal frequency spreading, is assumed by their product $\big(ds(t) = c(t) \cdot d(t)\big)$. It can be digitally realised by an inclusive-AND ($ds_i = d_i \oplus c_i$). The spreading factor (Ns) is therefore the ratio of the code rate with the symbol rate (Tc/T).

A sequential generator generates the PN-code. These sequences are used because of their particular autocorrelation functions: the correlation between two PN-codes is one if they are equal and aligned and is zero elsewhere [6]:

$$Rc(0) = \int_{-\infty}^{+\infty} c(t)c(t)dt = 1\,,$$

$$Rc_i c_j(\tau) = \int_{-\infty}^{+\infty} c_i(t)c_j(t-\tau)dt = 0\,,$$

$$Rc(\tau) = \int_{-\infty}^{+\infty} c(t)c(t-\tau)dt = 0 \quad \text{for} \quad \tau \neq 0\,.$$

Thanks to that property, the correlation between the received signal and the local PN-code allows the system synchronisation and also realises the processing gain (Gp). The processing gain increases the SNR at the input of the demodulator. There are several PN-code families. We chose the M-sequences because they are easily generated with linear feedback shift-registers. Moreover, they are used in many actual norms (especially UMTS) [7]. We have modelled two different types of generator as shown in Fig. 4; g_i represents some switches which are closed or opened according to the primitive polynomial $g(D)$. For example, the polynomial [2, 3] (010011 in octal) is written $g(D) = 1 + D + D^4$, and only the switches g_1 and g_4 are closed (g_0 being always closed).

The Galois generator is potentially faster, but the Fibonacci one's can be initialised with the PN-code loaded serially. This property is used in the RASE synchronisation.

At the receiver, one conception problem is to choose between intermediate frequency or baseband architecture. In a simple way, the BB is interesting because the signal processing is easier to implement (simple digital processing) and also because the signal (and system) bandwidth is two times smaller than IF one's. In the other way, the RF front/end is harder to realise because the frequency transposition goes to zero. This transposition to the baseband leads to conception problems as IIP2, LO leakage, self-mixing, components BF noise [8, 9]. Using IF architecture allows ignoring those problems. The drawback of this technique is the larger bandwidth, which leads to the realisation of chips with bigger size. This is a conception constraint because the power consumption and the price

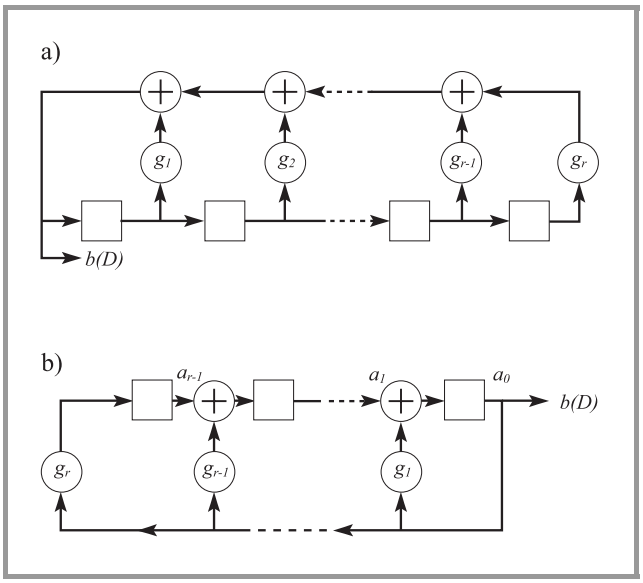


Fig. 4. PN-code generator models: (a) Fibonacci generator; (b) Galois generator.

increase with the chip size. The two types of architectures considered (Fig. 5) have been modelled (for the direct link elements), under SABER, in order to compare their performances. Also, because it uses the receiving signal, the tracking element has been also modelled for a BB and an IF architecture [10].

We can describe the three main elements as followed:

The I/Q demodulator [11] represents the RF front/end. Because the present study is focused on the conception of the signal processing stage, this I/Q demodulator does not take into account all the imperfections of this stage (as non-linearity or BF noise) but allows to consider the carrier impact on the system (filtering, harmonic imbalance...).

The despreader assumes the decorrelation of the PN-code from the received signal. As it despreads the signal, it realises the processing gain (Gp), that is to say the output despreader SNR is the input SNR plus Gp ($S/N_{out} = S/N_{in} + 10 \log Ns$) [12]. In the BB implementation, this function is assumed by a simple logic gate followed by an integrate and dump filter (I&D). The I&D filter sums all the chips in a symbol and dumps the value at the end of the integration. As shown in Fig. 5, for the IF architecture, the integration must be assumed by a filter (because of the carrier) and the product must be done by a multiplier (for which the digital implantation is heavy).

The differential demodulator is assumed by the product of the received symbol with the previous one [13]. Again this function is made by a logic gate and a latch in BB, and by a multiplier in IF.

2.3. Synchronisation elements

The synchronisation principle is shown in Fig. 6a. First, it is necessary to find the PN-code phase (origin) in the re-

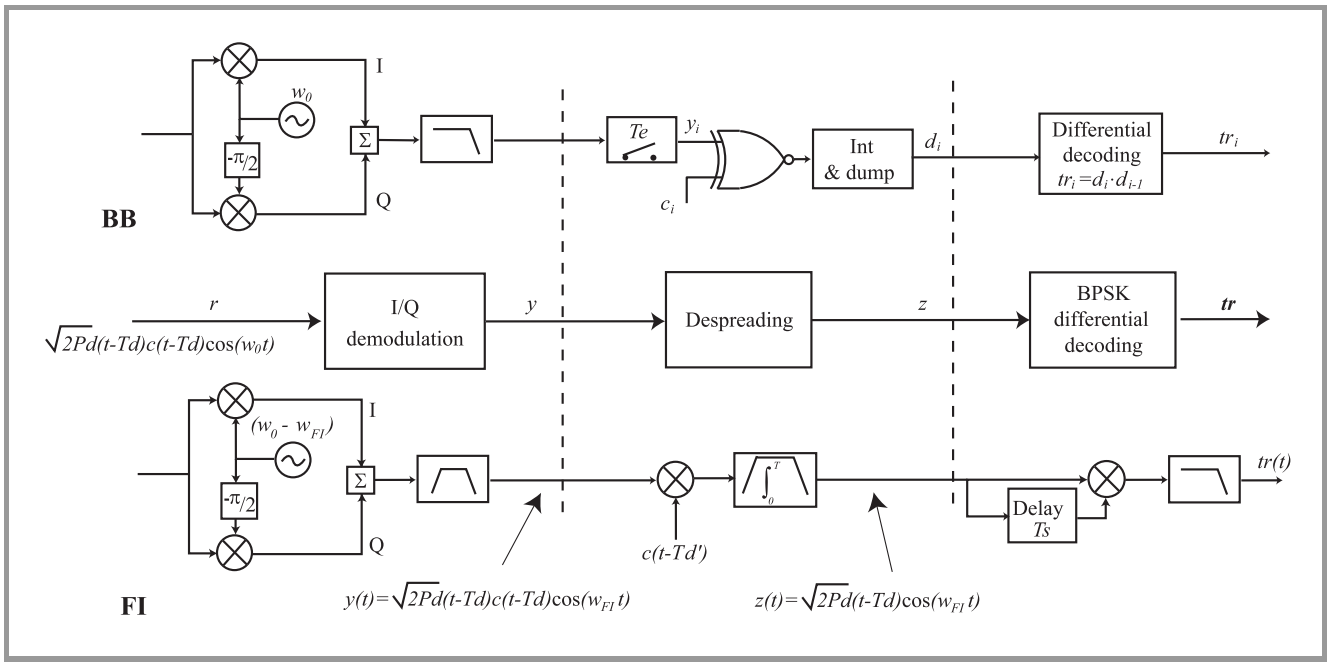


Fig. 5. Principle and models of IF and BB DBPSK receivers.

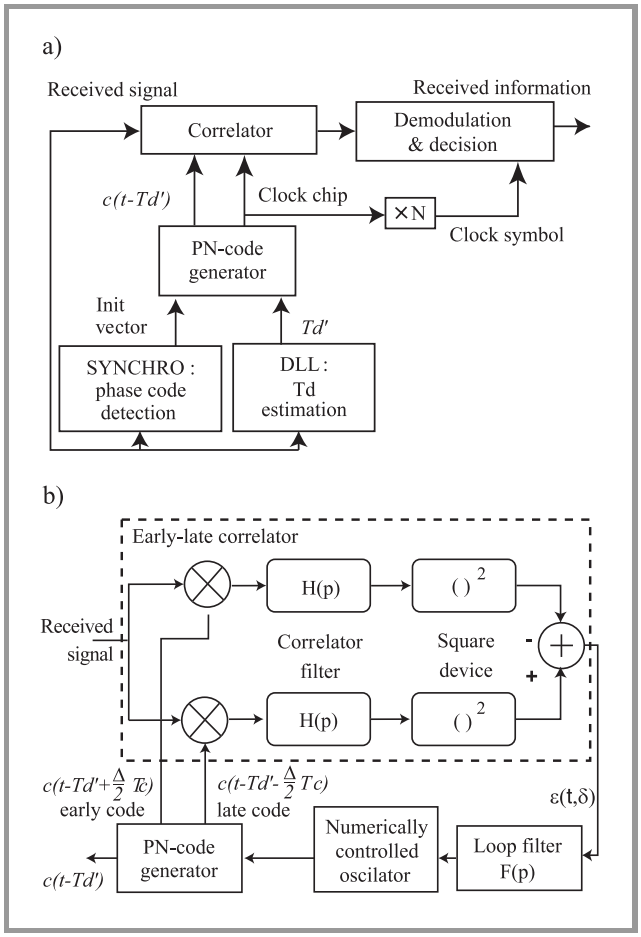


Fig. 6. (a) Synchronisation and acquisition principle; (b) DLL model.

ceived sequence. This function is the acquisition (or initial synchronisation). Next, because the delay in the physical channel is time-variant, it is necessary to evaluate continuously this delay. This function is the tracking.

The tracking is achieved by the DLL [14]. Its model is shown in Fig. 6b. The DLL evaluates the channel delay ($Td(t)$ being the delay in the channel and $Td'(t)$ its estimation) and generates a PN-code slaved by the delay error ($\delta = Td(t) - Td'(t)$). The early-late correlator output signal ε varies with δ thanks to the correlation between the received signal and the PN-code. In the BB architecture, the filter $H(p)$ and the square-low device $()^2$ are replaced by an I&D for the implementation of the early-late correlator [15].

The first acquisition system modelled on SABER is the rapid acquisition based on sequential estimation (RASE) [16], presented in Fig. 7a. Here the local PN-code is built in phase thanks to the estimation of the PN-code received.

Initially, the switch is on position 1. The estimation is achieved by sampling (in BB) or by an I/Q demodulation (in IF). Then, the PN-code estimated is used to load (serially) the $\log_2(N+1)$ shift registers of a Fibonacci generator. This load allows the generation of a PN-code synchronised with the one received. When the shift registers are loaded (after $Tc\log_2(N+1)$ second), the switch is put on position 2 and the generator operates normally.

Then the correlation between the received signal and the local PN-code is estimated. If the correlation occurs, the switch stays on position 2, in the other way round it means that a wrong estimation has been made and the switch is put again on position 1. The interest of this technique is the short acquisition time ($Tc\log_2(N+1)$ in the ideal case), but its drawback is its high noise sensitivity.

The other code acquisition technique implemented with the MAST language is the serial search presented in Fig. 7b.

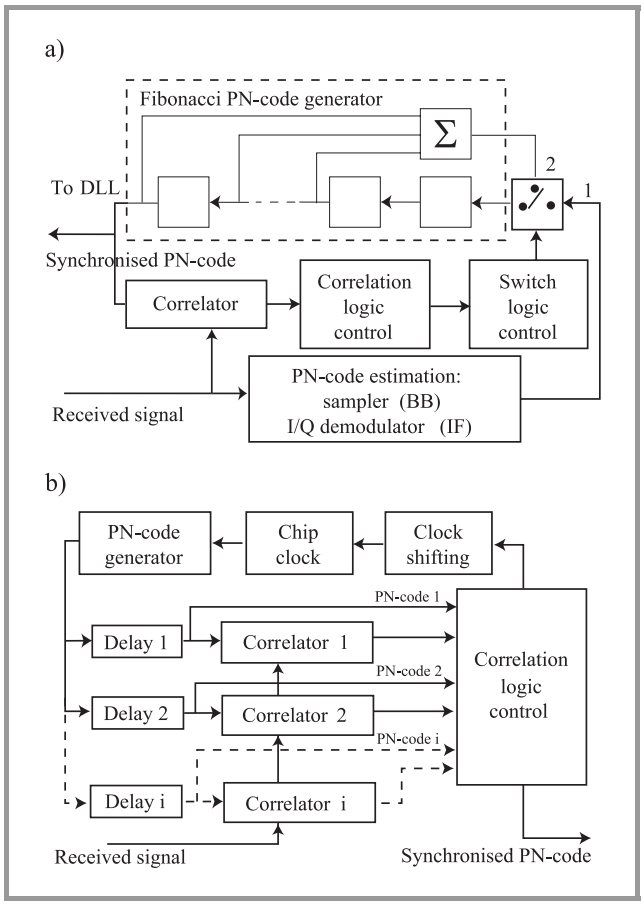


Fig. 7. (a) Rapid acquisition based on sequential estimation model; (b) serial search model.

It was developed by Sage [17]. The principle lies on the estimation of the synchronisation by the correlation of the received signal with the local PN-code. A control device measures the correlation. If the correlation does not occur, the local PN-code is one chip shifted until the correlation occurs [18]. This technique drawback is the maximum acquisition time, which is $N \cdot Tcor$ (where $Tcor$ is the correlation time that is around T) in the noiseless case. This time is reduced by addition of multiple dwells as shown in Fig. 7b. On the other hand, this technique is more immune from noise because it uses the code correlation instead of using the code estimation.

3. Simulations results

3.1. Modulation/demodulation

In Fig. 8a, the different architectures performances in an AWGN channel are plotted. The TEB estimator and the AWGN channel are modelled in MAST language. The despreader is suppressed in order to compare with theory. In Fig. 8a, the model 1 represents the IF architecture, the model 2, the BB one's and the model 3 is the same as model 2 with a matched filter (which is an I&D when the signal is BB) before the demodulator. It appears that model 3 matches with the theory. The divergence observed for the other models is due to the lake of matched-filter, which is necessary to insure the maximum likelihood criterion [DC]. This consideration is important because such filter is quite difficult to make when the signal is IF.

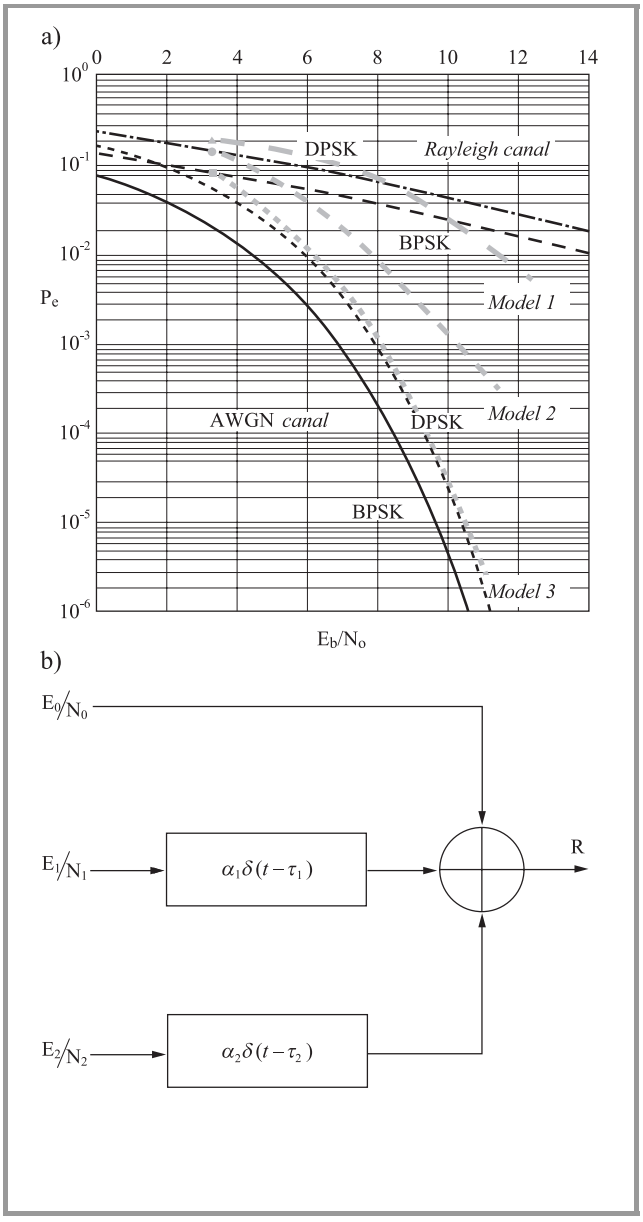


Fig. 8. (a) TEB of the 3 models implanted; (b) multi-path model.

3.2. Despreader

Several studies have been made. First, our virtual prototypes allow us to compare the Gp for the two architectures considered. For a PN-code length equal to 15, the BB processing gain obtained is 12.3 dB (in the ideal case where the sampling is done on an infinite bit number, which is the case the closest to the IF case). For the IF architecture, the processing gain is about 10.8 dB. This highlights the BB for the realisation of this function.

Next, two key parameters associated with the digital implantation are considered [15]. The first one is the sampling frequency, a key parameter when it comes to engineering a good performance-consumption trade-off during reception, and the second one is the number of bits needed at the analog to digital converter during reception that significantly impacts the circuit size. To perform this simulations, we studied the evolution of the signal to noise ratio at the correlator output versus signal to noise ratio at input for different values of sample frequency and number of bits of the analog to digital converter. We also read the process gain value for each simulation. The results are summarised in Table 1.

Table 1
Simulated process gain value for typical sample frequencies and receiver ADC bits number ($N = 15$)

Sampling frequency [MHz]	Process gain [dB]	Number of bits (DAC)	Process gain [dB]
80	8.1	3	11
120	10.54	4	14.3
160	11.2	6	12.8
		8	14.1
		infinite	12.3

A sampling frequency of 80 MHz is markedly inadequate to maintain a significant process gain. Conversely, a frequency higher than or equal to 120 MHz warrants a gain at least equal to 11 dB, close to the maximum theoretical value of 11.8 dB.
The number of bits of the DAC affects the size of the buses in the ASIC circuit and therefore the size and final consumption of the global chip. The study shows that 4 bits is a sufficient value to maintain an amply adequate process gain in the order of 14 dB. Conversely, for a number of bits equal to 3, one notes an important drop of 3 dB on the process gain relative to the maximum theoretical value.

Table 2
Additional SNR due to multi-path

Topology	Characteristics ($Tc = 25$ ns)	S/N out
One user: $E_0 = E_1 = E_2 = U$	$N_0 = N_1 = N_2 = 15$, $\tau_1 = 12.5$ ns, $\tau_2 = 40$ ns, $\alpha_1 = \alpha_2 = 1.1$	+13.5 dB
Two user: $E_0 = E_1 = U_1$, $E_2 = U_2$ (R is on U_1)	$N_0 = N_1 = 256$, $N_2 = 15$, $\tau_1 = 10$ ns, $\tau_2 = 40$ ns, $\alpha_1 = 0.75$, $\alpha_2 = 1$	+48 dB
Two user: $E_0 = E_1 = U_1$, $E_2 = U_2$ (R is on U_1)	$N_0 = N_1 = 15$, $N_2 = 256$, $\tau_1 = \tau_2 = 40$ ns, $\alpha_1 = 0.75$, $\alpha_2 = 1$	+8.5 dB

At last, we have studied the multi-path channel impact on the correlation thanks to the model shown in Fig. 8b which allows simulating real topologies.

The perturbations due to the different paths are measured in term of additional SNR (at the output despreader) comparatively to the case of a single path. The results are summarised in Table 2.
More than the trivial observation which is to say: long codes prevent from multi-path, this simulation shows that the paths arriving during the time chip ($\tau < Tc$) increase the SNR. In other terms, to take advantage of time diversity and DS-SS, the signal bandwidth must be around the coherence bandwidth. In the other way, especially if $1/Tc$ is bigger than the coherence bandwidth (which is around 10 MHz in our indoor application), a RAKE receiver is needed to use time diversity and so prevent from multi-path fading.

3.3. DLL

The DLL performance analysis is done thanks to tracking jitters variance σ_s [18]. The Fig. 9a shows a comparison between the theoretical and measured variance of the BB and IF model. First, it validates our MAST model, since the measure is close to the theory. Next, it shows that the BB is more immune to the white Gaussian noise thanks to its correlation filter. Indeed, thanks to our virtual prototype, we highlight the fact that the early-late correlator filter bandwidth can be set to $1/T$ in the BB case whereas it must be around $1/Tc$ in the IF architecture. In IF, a smaller bandwidth results in a loss of energy that degrades the DLL performance.
The Fig. 9b compares a DLL architecture where the tracking is done on a entire chip ($\Delta = 1$) with one where the tracking is done on a half chip ($\Delta = 0.5$). Here, the simulation result is quite interesting because the measure gives the $\Delta = 0.5$ architecture to be the best when the theory gives the $\Delta = 1$ architecture. However, the $\Delta = 0.5$ architecture is said to be better in practice [20], and this contradiction is explained by the divergence of the theoretical model when $\Delta \neq 1$. This shows how such virtual prototyping is helpful in system conception, especially to choose architecture.

3.4. RASE vs serial

Because the simulation of the code acquisition mean time is too long, (there are too many random parameters), we have studied the noise limit for which this time increases exponentially. The results presented in Fig. 10 compare the RASE and the serial search techniques for a BB signal. It appears that the serial search (even if it is slower than RASE) is more immune to the noise since its limit is around -8 dB against -4 dB for the RASE.

3.5. Simulation summary

This unified description with the MAST language was helpful for the conception and the design of our system. The previous study highlights the BB architectures for their simplicity and most of the time for their better performances (despreader, demodulator, DLL). For the synchronisation

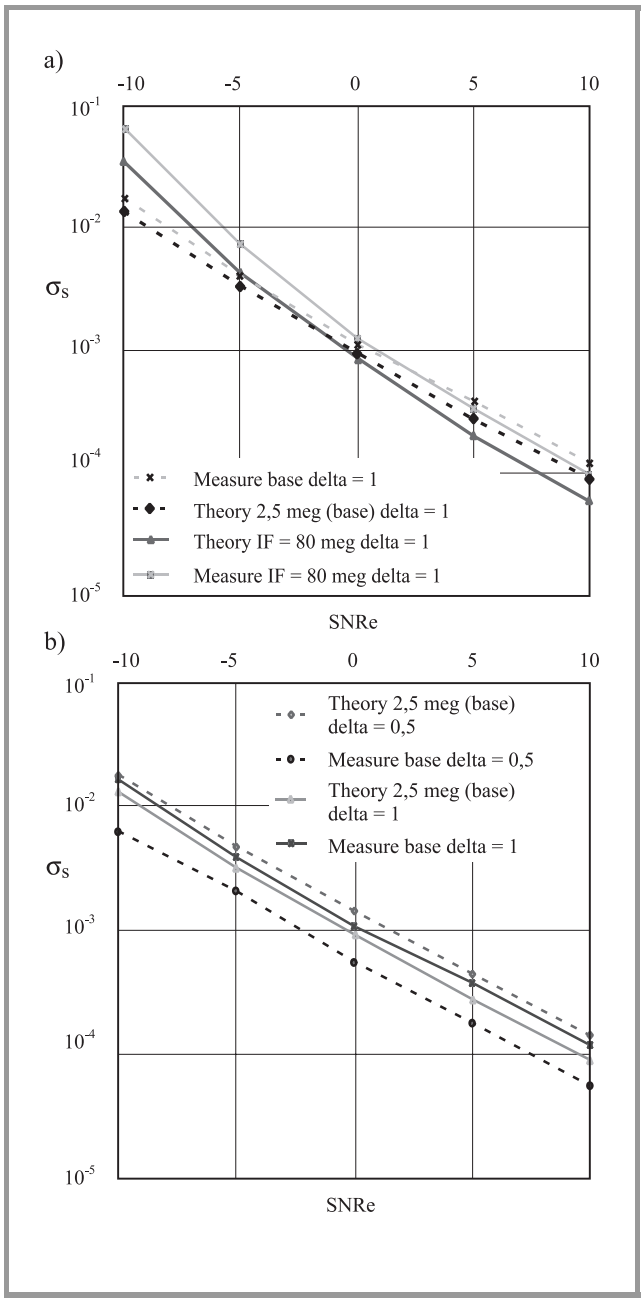


Fig. 9. DLL performance analysis: (a) δ_S for BB and IF architecture; (b) δ_S for different Δ architecture.

sub-system, the $\Delta = 0.5$ DLL and the serial search are retained according to there performances in noisy environment. In term of system constraints, our virtual prototype shows that a double integrator is needed in serial search. Moreover, the multi-path analysis shows that the system bandwidth must be < 10 MHz to take advantage of time diversity. At least a matched-filter is needed in receiver. In the BB architecture, the matched filtering is assumed by the correlator I&D filter placed just before the demodulator. In terms of parameters, the simulation shows that the BB DLL correlator integration time can be set to *Tc*. Also, for a digital implementation, the simulation shows that a frequency sampling of 160 MHz and a 4 bits DAC is needed to achieve the processing gain.

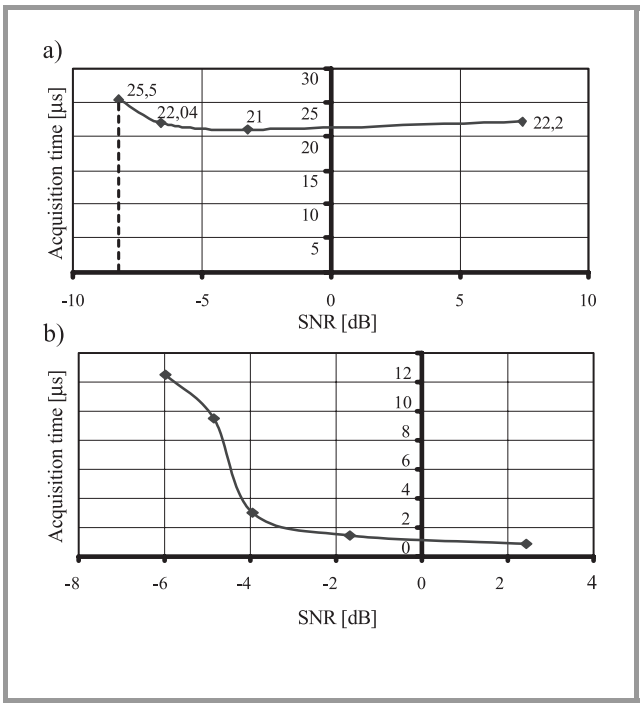


Fig. 10. (a) Serial search; (b) RASE.

4. ASIC conception

Since the architecture and the key parameters are defined and studied in the previous part, the integration of the digital signal processing functions on an ASIC circuit is now presented. First, the methodology and the baseband digital stage are presented. Next, some simulation results derived from the VHDL encoding of the main digital functions and the corresponding circuits are presented using CADENCE software. These simulations will allow us to quantify the performances reached (in terms of costs, consumption and surface area) of the ASIC circuit as a function of the technologies put forward (CMOS and BiCMOS).

According to the global functional system presented in Fig. 2, we present in the Table 3 the different numerical stages that we decided to implement on the ASIC for the complete transceiver. As seen in the Table 3, a particularity of our circuit is the possibility of transmitting and receiving at different rates from 19 kbps ($N = 2047$) to 5.71 Mbps ($N = 7$). So, a low rate corresponds to a high process gain, and therefore the channel noise perturbation can be reduced. Likewise, a higher rate induces a low process gain and the demodulation is more difficult to achieve.

4.1. Methodology

Our aim is to produce a low consumption and a low cost circuit. One of the original features of our study lies in the

Table 3
Numerical functions implanted in ASIC circuit

Transmitter		Receiver	
Differential encoder	yes	Differential decoder	yes
PN code generator	yes	PN code generator	yes
Multi-rate command ($N = 7$ to 2047)	yes	Multi-rate command ($N = 7$ to 2047)	yes
Spreading correlator	yes	Despreading correlator (integrate & dump)	yes
Nyquist filter	no	Nyquist filter	no
		DLL	no
		Serial search	no

development of a source code using a VHDL portable onto different technologies.

After validation of our system, we now investigate the implementation of main transmitter and receiver digital functions. To perform the simplest ASIC tests, we integrated a state machine type module in order to synchronise and adequately track the received signal.

The available technologies in our laboratory are A.M.S. type (Austrian Mikro Systeme), as indicated in Table 4. Under CADENCE environment, functional simulations are carried out with Affirma and placement-routing with silicon ensemble. Under SYNOPSYS, we perform the synthesis with design analyser. The overall design chart given in Fig. 11 describes the top-down approach starting from the VHDL specification.

Table 4
Available technologies (X)

AMS technologies	0.8 μm	0.6 μm	0.35 μm
CMOS	X	X	X
BiCMOS	X	–	–

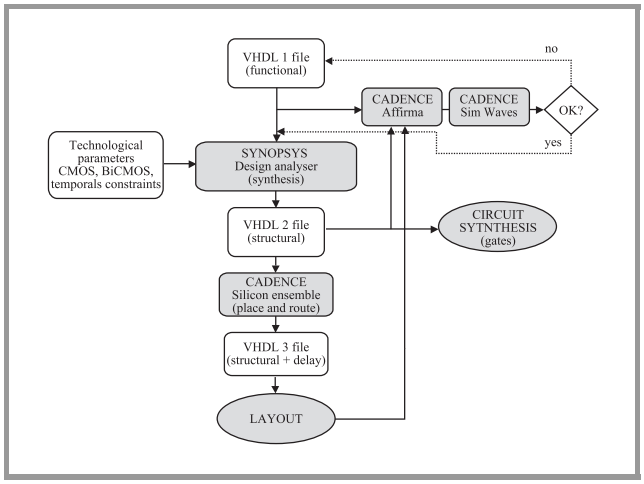


Fig. 11. Top-down ASIC circuit design under CADENCE and SYNOPSYS.

4.2. Transmitting and receiving circuits synthesis

Figure 12 shows the synthesised functions of the global transceiver, including the main digital functions as indicated in Table 1.

Concerning the transmitter, a Galois type PN code generator adjustable in length by means of a 4 bits word ($N_{\max} = 2047$) has been implemented. In the same way, a drive module allowing to generate the clock frequencies corresponding to the different possible data rates (19 kbps to 5.71 Mbps) and to synchronise the digital frame to be transmitted with the local PN code, has been implemented.

For the receiver, we chose an 8 bits word size for the analog-to-digital converter output in order to facilitate the first functional tests of the ASIC. Then, before the final production of the complete ASIC (including the synchronisation and tracking units of the PN code received), we will synthesise a 4 bits architecture (generic VHDL code).

First, we synthesised the transceiver using various technologies (see Table 5) and compared the area, power consumption and production cost of the ASIC using the AMS technology available in the lab.

The CMOS 0.8 μm technology is the least expensive. Nonetheless, for an equal cost, the 0.6 μm technology occupies twice as less surface and causes a 30 percent drop of the power consumption. With regard to the currently embedded functions, and in spite of a very important area and power consumption fall, a 0.35 μm technology seems too onerous for our application.

Finally, we choose a 0.6 μm CMOS technology, which is a good trade-off in terms of cost and performances. We can notice that interconnections use 60% of the total area of the transmitter chip. The circuit is actually in course of realisation and the first tests will be performed in few months.

5. Conclusion

Using a unified specification approach, we realised a virtual prototype of a complete DS-SS DBPSK transceiver under SABER software (using the MAST language). This stage conception allowed us to specify the optimum architecture taking into account the environment and the subsystems interactions. Baseband architecture is chosen in regard of its performances. Also a $\Delta = 0.5$ DLL and the serial search is chosen for the realisation of the synchronisation elements. Moreover, the functional models allow us to determine important parameters as the DLL correlator integration time ($1/T$ in BB and $1/Tc$ in IF) or the system bandwidth (< 10 MHz). Also, the impact of three key parameters related to the digital stages, i.e., sampling frequency ($f = 160$ MHz) and converter bus size (4 bits), is analysed.

Portable VHDL encoding the basic digital functions enabled us to quantify the system performances under CADENCE and SYNOPSYS software. Simulation results shown that interconnections and drive circuits oc-

Table 5
Sensitivity to various AMS technologies ($V_{DD} = 4.5$ V; $T = 350$ K)

Technologies	Area [mm^2]	Power consumption [mW]	Price/mm^2 [euro]	Total price [euro]	Minimal area [mm^2]
0.8 μm CMOS	2.16	55	220	475	4
0.8 μm BiCMOS	2.7	68	440	1188	3
0.6 μm CMOS	0.97	42	300	291	4
0.35 μm CMOS	0.32	4	490	156.8	12

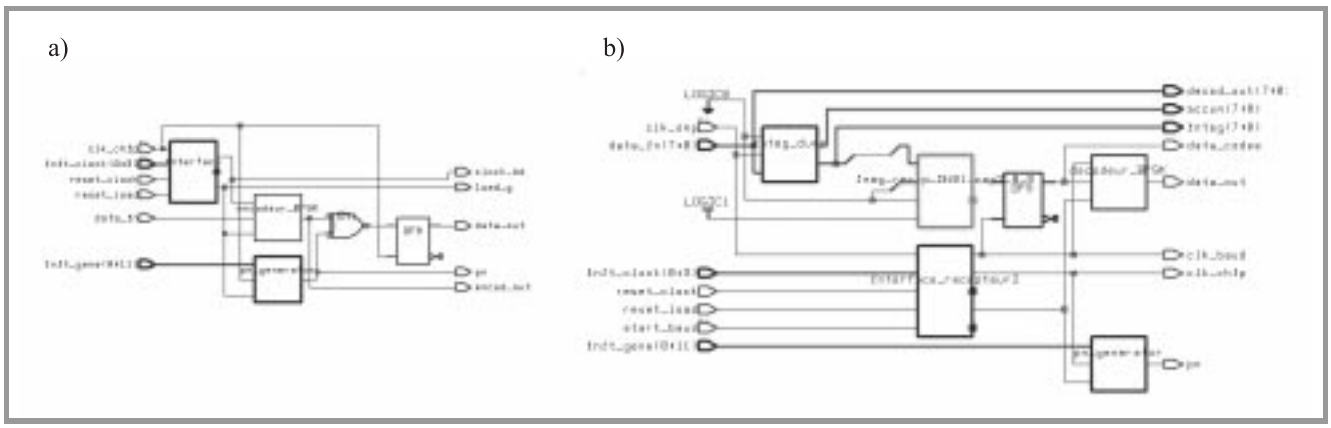


Fig. 12. Synthesised transmitter (a) and receiver (b) under design analyser.

cupy a substantial area compared to elementary functions. Finally, we have investigated the impact of CMOS and BiCMOS technologies and found a good trade-off in terms of cost and performance with a 0.6 μm CMOS technology.

All digital functions have not been embedded yet. The DLL and the Nyquist filter are being developed and precedent results could be perceptibly affected.

References

[1] R. Kohno, R. Meidan, and L. B. Milstein, "Spread spectrum access method for wireless communication", *IEEE Commun. Mag.*, pp. 58–67, 1995.

[2] R. A. Scholtz, "The origins of spread spectrum communications", *IEEE Trans. Commun.*, vol. 30, no. 9, pp. 822–853, 1982.

[3] J. G. Proakis, *Digital Communication*. Mac Graw Hill, 1995.

[4] H. M. Chang and M. H. Sunwoo, "Implementation of a DSSS modem ASIC chip for wireless LAN", in *Proc. IEEE Workshop Signal Proc. Syst. (SIPS'98)*, New York, 1998, pp. 243–252.

[5] D. C. Kemdirim and J. S. Wight, "DS SSMA with some IC realisation", *IEEE J. Solid State Circ.*, vol. 8, pp. 663–673, 1990.

[6] R. L. Peterson, R. E. Ziemer, and D. E. Borth, *Introduction to Spread Spectrum Communications*. New York: Prentice Hall, 1995.

[7] P. W. Baier, P. Jung, and A. Klein, "Taking the challenge of multiple access for third-generation cellular mobile radio systems – an European view", *IEEE Commun. Mag.*, pp. 82–89, 1966.

[8] A. Abidi, "Direct conversion radio transceivers for digital communication", *IEEE J. Solid State Circ.*, vol. 30, no. 12, pp. 1399–1410, 1995.

[9] B. Razavi, "Design consideration for direct conversion receiver", *IEEE Trans. Circ. Syst.–II*, vol. 44, no. 6, pp. 428–435, 1997.

[10] S. Bourdel, E. Campo, L. Andrieux, and P. Melet, "Modelling and comparative trial of baseband and IF DLL in a noisy environment", in *Proc. IEEE Sixth Int. Symp. Spread Spectr. Techn. Appl. (ISSSTA 2000)*, Parsippany, USA, 2000, pp. 685–689.

[11] T. Noguchi, Y. Daido, and J. A. Nossek, "Modulation techniques for microwave digital radio", *IEEE Commun. Mag.*, vol. 24, no. 10, pp. 21–30, 1986.

[12] R. L. Pickholtz, D. L. Schilling, and L. B. Milstein, "Theory of spread-spectrum communications – a tutorial", *IEEE Trans. Commun.*, vol. 30, pp. 855–884, 1982.

[13] E. Fernandez and M. Mathieu, *Les Faisceaux Hertziens Analogique et Numériques*. Paris: Dunod, Bordas and CENT-ENST, 1991.

[14] M. K. Simon, "Noncoherent pseudo-noise code tracking performance of spread spectrum receiver", *IEEE Trans. Commun.*, vol. 25, no. 3, pp. 327–345, 1977.

[15] B-Y. Chung, C. Chien, H. Samueli, and R. Jain, "Performance analysis of an all-digital BPSK direct sequence spread-spectrum IF receiver architecture", *IEEE J. Select. Areas Commun.*, vol. 11, no. 7, pp. 1096–1107, 1993.

[16] R. B. Ward and K. P. Yiu, "Acquisition of pseudo-noise signals by recursion-aided sequential estimation", *IEEE Trans. Commun.*, vol. 25, no. 8, pp. 784–794, 1978.

[17] G. F. Sage, "Serial synchronisation of pseudo-noise systems", *IEEE Trans. Commun.*, vol. 12, pp. 123–127, 1964.

[18] A. Polydoros and C. L. Weber, "A unified approach to serial search spread spectrum code acquisition". Part I. "General theory", *IEEE Trans. Commun.*, vol. 23, pp. 550–560, 1984.

[19] J. K. Holmes and L. Biederman, "Delay lock loop mean time to loose lock", *IEEE Trans. Commun.*, vol. 26, no. 11, pp. 1549–1556, 1978.

[20] W. Sheen and G. L. Stüber, "Effects of multipath fading on delay-locked loops for spread spectrum systems", *IEEE Trans. Commun.*, vol. 42, no. 2/3/4, pp. 1947–1956, 1994.

Sylvain Bourdel, received the Ph.D. degree in electronics from the National Institute of Applied Sciences of Toulouse, France in October 2000. Specialised in microelectronics, he is working in ICARE research team and LAAS/CNRS laboratory of Toulouse, on high-level language description applied to specifications for telecommunication's microsystems using specially spread spectrum techniques.
e-mail: bourdel@laas.fr
Groupe ICARE de l'IUT B Blagnac
Université Toulouse II
1 Place Georges Brassens – BP
73-31703 Blagnac, France
Laboratoire d'Analyse et d'Architecture
des Systèmes du CNRS
7 Avenue du Colonel Roche
31077 Toulouse, France

Eric Campo, received the Ph.D. degree in electronics from the National Institute of Applied Sciences, Toulouse, France in 1993. Now, he is an Assistant Professor at the Technology University Institute of Blagnac. Since 1994, he is with Microsystems and Systems Integration group of the LAAS Laboratory and with ICARE team of IUT. His work is focused on the home automation field: multisensor monitoring applied to the elderly, definition of new functions at home, functional modelling of CDMA spread spectrum systems and realisation of wireless communication systems.
e-mail: campo@laas.fr
Groupe ICARE de l'IUT B Blagnac
Université Toulouse II
1 Place Georges Brassens – BP
73-31703 Blagnac, France
Laboratoire d'Analyse et d'Architecture
des Systèmes du CNRS
7 Avenue du Colonel Roche
31077 Toulouse, France

Patrick Melet, received the post-graduate diploma in conception of microelectronic and microsystem from Paul Sabatier University, Toulouse, France in 1997. Now, he is working toward the Ph.D. degree in conception and evaluation of a spread spectrum transceiver ASIC for home automation, in ICARE research team and LAAS/CNRS laboratory of Toulouse. His primary research interests include architecture design of digital telecommunication and high-speed low-power ASIC design.
e-mail: melet@laas.fr
Groupe ICARE de l'IUT B Blagnac
Université Toulouse II
1 Place Georges Brassens – BP
73-31703 Blagnac, France
Laboratoire d'Analyse et d'Architecture
des Systèmes du CNRS
7 Avenue du Colonel Roche
31077 Toulouse, France

Laurent Andrieux, received the Ph.D. degree in electrical engineering from the University Paul Sabatier of Toulouse, France, in 1995. His thesis was devoted to high power GaAs heterojunction bipolar transistor in S band for mobile telecommunications. Since 1996, he has been an Associate Professor in the Telecommunications and Networks Department of the Technology University Institute of Blagnac. He is currently working in ICARE research team and LAAS/CNRS laboratory of Toulouse. His research interests include the modelisation and synthesis of global CDMA spread spectrum systems in noisy environments. He specially focuses his works on the development of the ASIC and radio frequency stages of a demonstration model for wireless indoor communications.
e-mail: andrieux@laas.fr
Groupe ICARE de l'IUT B Blagnac
Université Toulouse II
1 Place Georges Brassens – BP
73-31703 Blagnac, France
Laboratoire d'Analyse et d'Architecture
des Systèmes du CNRS
7 Avenue du Colonel Roche
31077 Toulouse, France

Paper

Simulating capture behaviour in 802.11 radio modems

Christopher Ware, Tadeusz A. Wysocki, and Joe Chicharo

Abstract — Simulation plays an important role in the performance evaluation of MAC protocols. Building simulation models which are able to accurately model physical behaviour is fundamental to the outcome of such techniques. Through both qualitative and quantitative comparison of experimental trace data against simulation results obtained using delay, power, and hybrid capture models, this paper investigates the performance of various packet capture models in the simulation analysis of the 802.11 PHY and MAC layer protocols. We illustrate these models are unable to accurately describe the fairness properties of the experimental data. A new model is proposed, Message Retraining, to describe the operation of an 802.11 receiver. We illustrate that the Message Retraining reception model is able to model the fairness characteristics obtained with an IEEE 802.11 radio modem more accurately than the previous capture models.

Keywords — wireless local area networks, ad hoc networks, capture models, medium access control protocol, fairness.

1. Introduction

The IEEE 802.11 wireless physical layer (PHY) and medium access control (MAC) protocols have resulted in the widespread adoption of local wireless area networking in recent years. However, recent experiment [1] has indicated that in many conditions, the potential exists for significant unfairness at the MAC layer. In this paper, we investigate the ability of capture models presented in literature [2 – 4] to provide a realistic representation of an IEEE 802.11 radio modem by undertaking both a qualitative and quantitative comparison with experimental data. Our results, combined with work investigating the impact of multiple access interference [5, 6] and parallel receiver structures [7], motivates the development of a new reception model. Thereafter, we will refer to this model as the Message Retraining capture model or simply Message Retraining.

The Message Retraining model allows the modem to retrain onto a newly detected signal [8] with a higher signal power, as a means of explaining the experimental data presented in [1]. Using simulation techniques, we illustrate that this model provides a very accurate description of the physical behaviour, and can be used to investigate the impact this has on higher layer protocols. Using two fairness metrics, we compare the performance of each of the capture models against measured experimental data. The major contribution of this paper is the qualification of the important role modem receiver behaviour plays in the operation of the higher layer protocols in varying signal conditions, and the inclusion of fairness as significant factor in the application of a given packet capture model.

We use an intuitive definition of **fairness** in this paper. Hosts should be able to achieve relatively equal transmission rates, and no host should be able to prevent others from gaining access to the channel for a sustained period. The network model considered in this paper is one involving hidden terminals over a semi-slotted 802.11 MAC/PHY layer. All nodes employ a common spreading code with no power control.

Capture can be considered to occur at two levels:

- **Modem capture** is a property of the radio modem and the modulation techniques employed [9]. Modem capture results in a given transmission being "captured" by the receiver while rejecting interfering frames as noise. Several models based on either power, time of arrival, or both, [2] have been proposed to evaluate the probability of a frame being captured by a receiver as a function of the number of interfering frames.

- **Channel capture** is induced by protocol timing, and results in a channel being monopolised by a single node, or subset of nodes in a given geographic region. Channel capture has been identified as a significant problem for multihop packet networks in many scenarios where disconnected topologies exist [10, 11], or higher layer retransmission and backoff timers are employed [12 – 14].

The original IEEE 802.11 standard [15] defines a medium access control protocol, and three distinct physical layers: an infra-red physical layer (IR), and two spread spectrum layers, one based on frequency hopping spread spectrum (FHSS), and another using direct sequence spread spectrum (DSSS). The 802.11 standard was updated [16] with the addition of the high rate (HR) physical layer extensions. This allowed the DSSS physical layer to operate at 5.5 Mbit/s and 11 Mbit/s in addition to the original 1 and 2 Mbit/s. Further extensions in the 5 GHz band employing orthogonal frequency division multiplexing have also recently been developed [17]. At the MAC layer, the distributed co-ordinate function (DCF) implements CSMA/CA, with an optional request-to-send/clear-to-send (RTS/CTS) handshake prior to transmission of DATA frames. Immediate positive acknowledgement is employed. This scheme is able to operate in a peer-to-peer ad hoc mode, being a fully distributed MAC protocol.

The remainder of this paper is organised as follows: Section 2 reviews the experimental results motivating this

work [1], Section 3 presents details of current capture models, as well as our Message Retraining reception model. Section 4 outlines our simulation results. Section 5 presents an investigation of the fairness properties of the simulation and experimental trace data, while Section 6 concludes the paper.

2. Review of experimental results

Recently published experimental results investigating the performance of the IEEE 802.11 MAC protocol in a hidden terminal topology [1], have illustrated that signal strength is a significant factor in determining which node is able to access the radio channel. The experiments uncovered a reliable and repeatable channel capture effect, in which a host with a higher received signal power (measured at the common node) was able to capture the channel.
Each trial involved two simultaneous 500 kbyte file transfers from both hidden senders into the common node. The linear topology with each end node being mutually out of range is illustrated in Fig. 1. A software package called **tcpdump** [18], is used at the central node to trace the progress of each file transfer. An 802.11 compliant wireless network interface is used in the ad hoc mode, employing an RTS/CTS handshake governed by the **aRTSThreshold** management information base (MIB) parameter. This parameter indicates the frame size, above which, an RTS/CTS exchange is initiated.

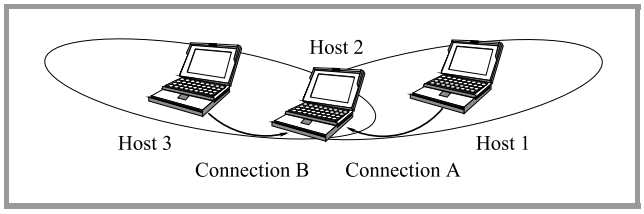


Fig. 1. Experimental topology – host 1 and 3 hidden terminals.

Several parameter combinations and signal to noise ratio (SNR) scenarios were investigated:

- **Trial 1.** No RTS/CTS equal power hidden terminal, SNR = 25 dB.
- **Trial 2.** 500 byte RTS/CTS equal power hidden terminal, SNR = 25 dB.
- **Trial 3.** 500 byte RTS/CTS near(25 dB) – far (20 dB) hidden terminal.
- **Trial 4.** 500 byte RTS/CTS controlled SNR.

The significant result from these experiments is a strong signal power dependence in the channel capture behaviour observed in each trial. In trial 1 without the RTS/CTS handshake and with equal signal power on the contending connections, random channel capture was exhibited. This result is expected, [12] as the combination of MAC and TCP backoff timers is known to result in a channel capture state for one of the contending nodes when hidden terminals are present. In trial 2, employing the RTS/CTS handshake, there was effective sharing of the channel when the signal power was equal on the contending connections. In trial 3, illustrated in Fig. 2 with a signal power differential of 5 dBm, the stronger connection was able to capture the channel, locking out the weaker contending host until the file transfer was complete. This is a reliable and repeatable observation. This result indicates that an unequal signal power scenario prevents the RTS/CTS handshake from providing fair access to the channel.

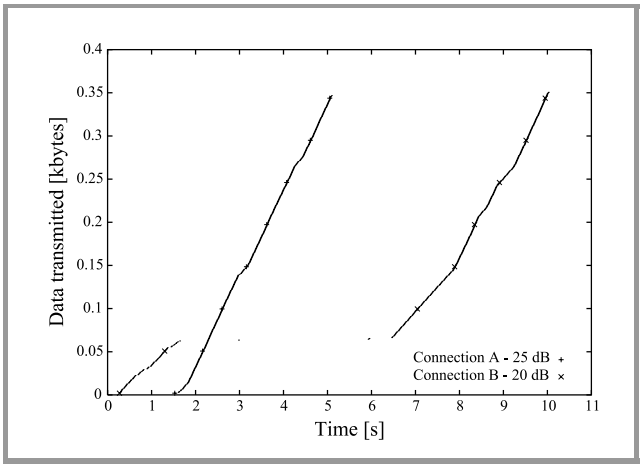


Fig. 2. Stationary SNR, RTS/CTS enabled.

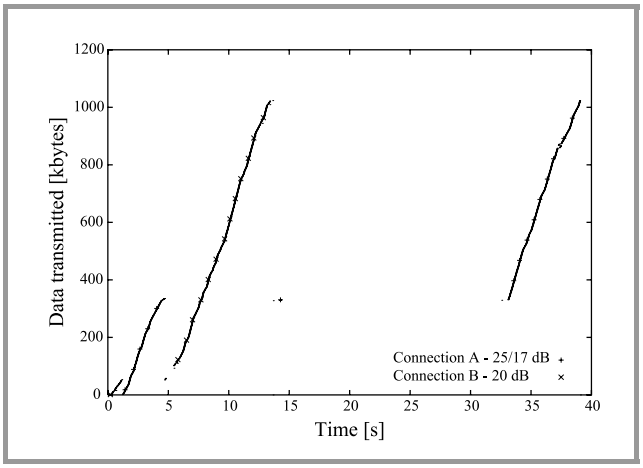


Fig. 3. Controlled SNR, RTS/CTS enabled.

To examine the signal strength dependence experimentally, a further trial was undertaken in which the signal power on one connection was controlled throughout the experiment. In this case, illustrated in Fig. 3, the SNR on connection A was controlled during the transfer. Connections A and B commence the test with a SNR of 25 dB and 20 dB respectively. Five seconds into the trial the signal power of connection A is reduced by 8 dBm through to the end of the experiment. The trace in Fig. 3, clearly illustrates

behaviour where the new stronger host, connection B, manages to "re-capture" the channel once the signal power of connection A is reduced.
Once connection B has finished, connection A is able to regain access to the channel. In each case, the connection which manages to capture the channel suffers relatively few TCP timeouts, and retransmissions are simply handled by the MAC and TCP retransmission mechanisms. Conversely, the contending connection will undergo continual timeout and exponential backoff at both the MAC and TCP levels. This results in significant unfairness under heavy load conditions. In the following sections we investigate the ability of capture models presented in the literature in describing this behaviour.

3. Capture models

The development of models describing the initial capture of a frame by a radio modem represents a significant body of literature [2 – 4]. The common goal of each model is to determine the probability with which a given frame may be captured by the receiver, as a function of the number of active stations, and the resulting channel throughput achieved.
There are two significant stages in the successful reception of a frame by a radio modem. Initially, the frame must be successfully detected and subsequently captured by the receiver. Following this, successful reception of the frame must be achieved in the presence of interference, from other transmissions and external noise sources. Most literature [2, 3] has considered the probability with which successful detection and capture of a frame at the start of a transmission slot occurs. The second aspect requires an understanding of the impact multiple access interference will have on the captured frame [5, 6, 9] and depends significantly on the modulation technique and spreading codes employed.
Capture models are often used when simulating the performance of wireless networks. The results presented in Section 2 however, suggest a more complex capture behaviour resulting in the significant unfairness evident in the traces. Further, in cases where hidden nodes are likely (e.g. a mobile ad hoc network) there is a strong possibility of late starting transmissions colliding with other signals at the common receiver. In a scenario where all nodes are able to sense carrier, slot boundaries are easily identified and defined, thereby reducing significantly the probability of a new transmission interfering with an ongoing transmission.
In scenarios where carrier sense mechanisms are unreliable, it is possible for a node to have no knowledge of an ongoing hidden transmission. This introduces the potential for an interfering transmission to arrive at a common receiver at any time during a slot. As illustrated in Fig. 4, this can be due to differences in the slot time boundaries observed by both hidden nodes. This is further complicated by the slot timing mechanisms within 802.11. Rigid slot boundaries are not maintained, requiring nodes to infer "slot" boundaries from the beginning and end of surrounding transmissions. Data transmissions are able to occupy multiple "slot times". Guard times are inserted between sensing an idle channel and transmitting (the distributed co-ordinate function inter-frame space, DIFS), or returning management frames (the short inter-frame space, SIFS) to maintain the semi-slotted channel. However, the lack of carrier from an opposing hidden node increases the possibility that the node will transmit at what appear random times to the common node.

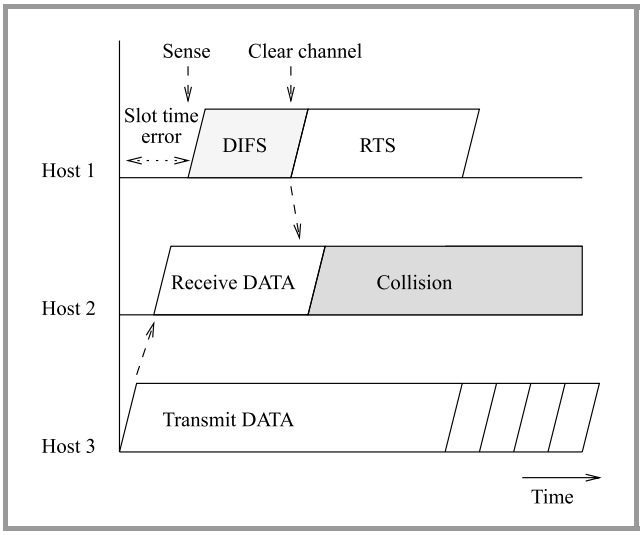


Fig. 4. Potential slot time error.

In Fig. 4, host 3 has commenced a data transfer prior to host 1 (being hidden from host 3) commencing a carrier sense operation. On sensing a clear channel, host 1 defers for a DIFS then transmits an RTS message. This collides with the data frame from host 3, illustrating the potential for a late starting transmission to interfere with an ongoing transmission.
In the following sections we briefly review the significant capture models considered in literature, with the Message Retraining reception model outlined in Section 3.4.

3.1. Delay capture

Delay capture originally described by Davis and Gronemeyer [3], enables the capture of a frame in a given timeslot, provided no other frame arrives within a given capture time, T_c of the initial frame. Only the initial frame is able to be received. Frame arrivals are assumed uniformly distributed on the interval $[0, T_u]$. The initial frame arrives at time T_1, and may be captured by the receiver provided that $T_i > T_1 + T_c$, where T_i is the arrival time of the ith frame. This model is chiefly controlled by the parameter T_c, governing the period of time required by a receiver to detect, correlate with, and lock onto the received signal. The larger the T_c/T_u ratio, the less effective the modem is at capturing a frame.

3.2. Power capture

Power capture, originally described with Rayleigh fading, and constant transmitter power [4], is described by the following inequality over the interval $[0, T_c]$:

$$P_{max} > \gamma \sum_{i=1}^{N} P_i \,, \tag{1}$$

where P_{max} is the power of the strongest of N signals arriving, each with power P_i, within the capture time T_c. The model allows a frame to be captured provided P_{max} is greater than the sum of the power of all other received frames, P_i, times the capture ratio, γ. The received signals are assumed to have phase terms varying quickly enough to allow incoherent addition of the received power of each frame. This model is the most commonly employed in the simulation of radio modems, allowing the first arriving frame in a slot to be received provided no other frame arrives within the capture time, T_c having a power violating (1). In the case where (1) is violated, no frame is captured.

3.3. Hybrid capture

The hybrid model was originally proposed by Cheun and Kim [2]. The power capture effect is used to increase the capture probability of the first arriving frame in a given timeslot, even though the delay model would otherwise indicate capture has not occurred. Capture occurs when the following inequality holds:

$$\gamma \sum_{i=2}^{N} P_i \left[T_1 + T_c - T_i \right] < T_c P_1 \,. \tag{2}$$

The total accumulated energy must be less than the energy received from the first packet, P_1 over the capture interval. This model results in a greater capture probability, reflecting the ability of a direct sequence spread spectrum receiver to correlate with the initially detected frame and reject other transmissions as noise.

3.4. Message Retraining reception model

Contrary to each of the models presented above, [8] describes an enhanced capture technique which allows a modem to successfully receive a signal that would otherwise be considered lost by the previous models. The modem implements a *Message in Message* process, whose function is to monitor the energy received on either antenna during reception of a frame. If an increase in energy beyond a given threshold, γ_{MR} is observed, the modem attempts to synchronise with and demodulate the new energy as a potential new signal. If this is achieved a retraining process allows the modem to prepare to receive this new frame once the prior transmission has finished.

This ability implies that each of the capture models previously described will result in a pessimistic capture probability for a frame over a given duration. The Message Retraining ability of the modem also extends the time scale over which capture must be considered. Retraining may take place at any time during frame reception, as opposed to the delay, power and hybrid capture models which consider a short duration at the start of a frame. We therefore propose an extended capture model, termed Message Retraining which incorporates the enhanced capture ability.

The model allows the modem to receive a new transmission (signal 2 in Fig. 5) which may arrive at a random time during the reception of a previous frame (signal 1 in Fig. 5), provided the new transmission has sufficient relative power to enable successful synchronisation and demodulation of the frame preamble. As indicated in [5], energy associated with the new transmission will have a significant impact on the BER observed at the correlator output for the original frame. Results presented in [5] indicate the previous frame will be unintelligible if the signal power difference between the new and existing transmission is greater than a threshold of $3-5$ dB. The Message Retraining model accounts for this by dropping the initial frame if a new frame is detected with a signal power greater than the current by the Message Retraining threshold, γ_{MR}. Successful reception of a frame, F_j will occur provided that over the duration of this transmission:

$$\gamma_{MR} \sum_{i=1, i \neq j}^{N} P_i < P_j \,. \tag{3}$$

This model allows for the successful reception of the strongest arriving frame received throughput its own duration, i.e. F_j will be successfully received provided no other frame arrives over the duration of F_j with a power greater than $P_j + \gamma_{MR}$ (measured in dBm). Furthermore, the initial frame may be successfully received provided the standard power capture equation, Eq. (1) holds.

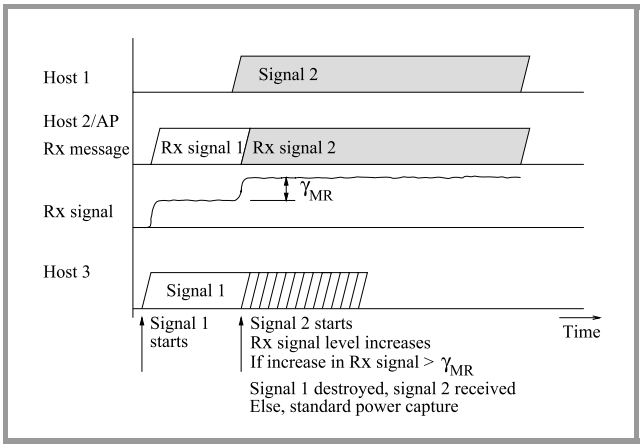


Fig. 5. Illustration of the operation of the Message Retraining model.

As the purpose of this paper is the evaluation of capture models for simulation, a more detailed analytic study of this model in terms of probability of successful reception of a frame is considered in future studies.

4. Simulation results

4.1. Simulation description

Each of the capture models described previously has been implemented using the *ns* simulation package. This package contains an 802.11 PHY/MAC layer model, as well as providing excellent implementations of higher layer protocols such TCP/IP, UDP, FTP etc. The channel model employed is an additive white Gaussian noise (AWGN) two-ray ground model. Capture decisions are made within each modem based on the received signal strength, capture threshold, and other relevant parameters for each model. Each node receives a copy of the transmitted packet and based on the received power, determines whether the transmission was observable or not. If the frame is received with sufficient power, a capture decision in accordance with each model is made prior to passing the frame up to the MAC protocol.

Table 1
Modem simulation parameters

Parameter	Value
γ	5 dB
P_t (nominal)	15 dBm
R_b	2 Mbit/s
Sensitivity	-95 dBm
f	2.412 GHz
T_c	120 μs

Parameters for the modem are listed in Table 1. The capture threshold is selected based on measurements presented in [1] and design parameters of the Message Retraining process in an 802.11 modem [8]. P_t represents the nominal transmitter power of the radio modem, R_b the channel bit rate (determined by the combination of spreading sequence and modulation technique employed), f the operating frequency, and T_c the capture interval which corresponds to the duration of the preamble and sync bits in the 802.11 PHY header. As the 802.11 standard requires that the PHY preamble and header are transmitted at 1 Mbit/s with an 11 chip Barker code using DBPSK modulation, or possibly 2 Mbit/s with the 11 chip Barker code using DQPSK modulation (where short the PLCP preamble/header option is available), we use a value of R_B at 2 Mbit/s.
Simulation trials of the controlled SNR experimental trial outlined in Section 2 were performed using each of the capture models. Each trial involves two hidden connections, as illustrated in Fig. 1 transferring data to a common node, using TCP. Connection B starts at 4 seconds, with connection A starting at 5 seconds. A total of 2048 packets of 512 bytes are transmitted over each connection. Initially, connection A is stronger than connection B by at least the capture ratio. At 10 seconds, this situation is then reversed for the remainder of the transfer.
In this section we present simulation traces of the controlled SNR experiment described in Section 2. The traces in Figs. 6 – 10 provide a qualitative means of comparing each of the capture models with recorded data. In Section 5 we undertake a quantitative comparison of each model with recorded data using fairness metrics over the length of the trace.

4.2. No capture

In the case where no modem capture is implemented, any colliding transmission at the common receiver will result in both frames being destroyed. Backoff and retransmission then results in reasonably effective sharing of the radio channel. In Fig. 6, alternating periods where either connection is able to dominate the radio resource are due to protocol timing interactions between MAC and TCP retransmission timers [12].

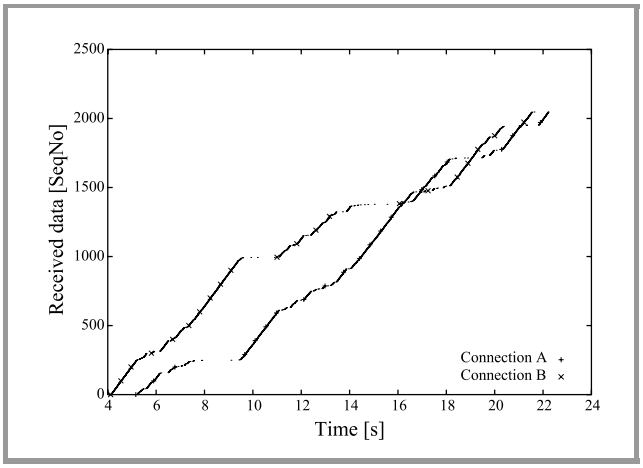


Fig. 6. No capture model – controlled SNR.

4.3. Delay capture

The delay capture model makes no account of signal strength characteristics. Figure 7 exhibits random periods during which one of the connections is able to capture the majority of the channel resource. This is again due to the interaction between MAC backoff timers and the TCP timers at the transport layer. Connection B also gains a slightly higher transfer rate than connection A, showing no evidence of the changed signal power at 10 seconds. This is due to connection B starting before connection A, and therefore having a larger TCP window at the time connection A commences. Connection B is able to expand it's TCP window without contention for the channel, whereas connection A must contend from the establishment of the TCP connection. The connection start times were staggered in this manner to match the experimental data in Section 2.

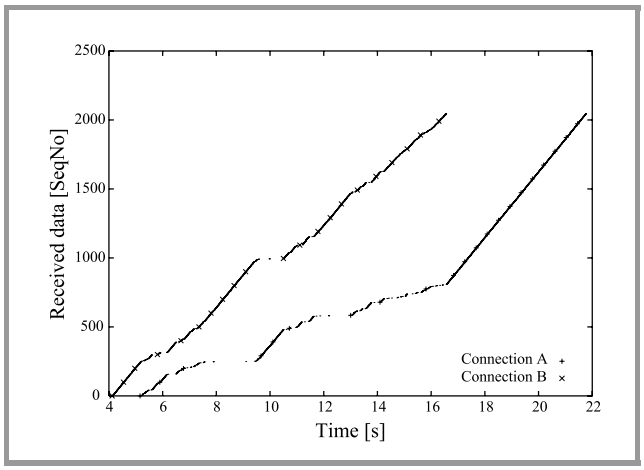


Fig. 7. Delay model – controlled SNR.

4.4. Power capture

The power capture model trace in Fig. 8 displays similar behaviour to the delay capture model. Neither connection is able to dominate. There is no evidence of the sustained channel capture exhibited in Fig. 3, nor any evidence of the transmission power change at 10 seconds.

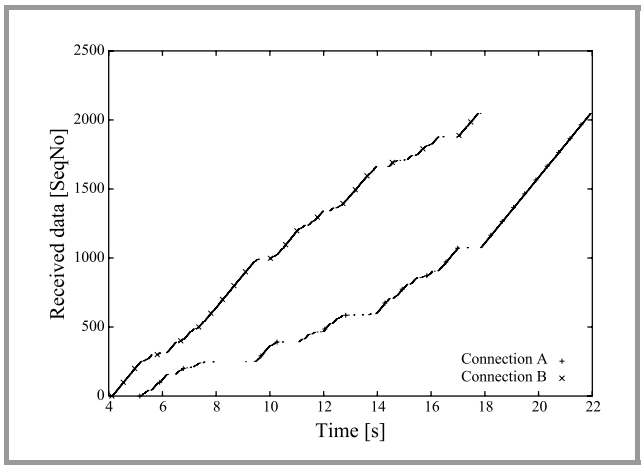


Fig. 8. Power model – controlled SNR.

4.5. Hybrid capture

Connection B again gains the advantage of a larger TCP window at the time connection A commences. As with the power model, there is no evidence of behaviour approaching that observed in the experimental trial.

In each of Figs. 6 – 9, the change in signal strength at 10 seconds has little impact on the channel access achieved by each connection. This represents a significant shortcoming for the delay, power and hybrid capture models, failing to reflect the impact varying signal strength characteristics have on connection quality.

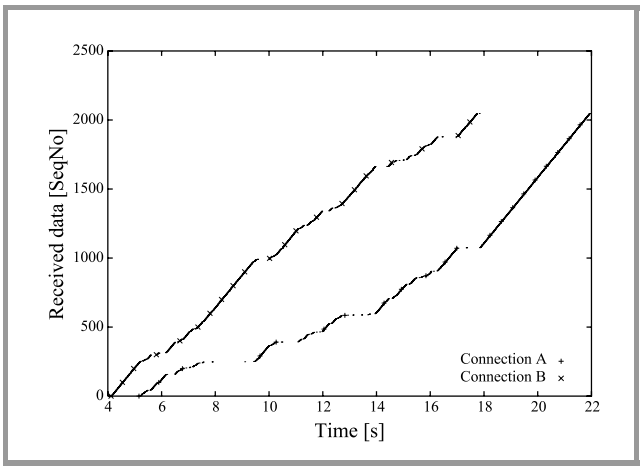


Fig. 9. Hybrid model – controlled SNR.

4.6. Message Retraining

The trace for the Message Retraining model in Fig. 10 appears to match the measured data of Fig. 3 quite closely. Once connection A commences as the stronger connection, connection B is prevented from gaining reliable channel access. 10 seconds into the trace, connection B, the new stronger connection, is able to capture the channel from connection A, which is in turn prevented from gaining fair access until connection B finishes.

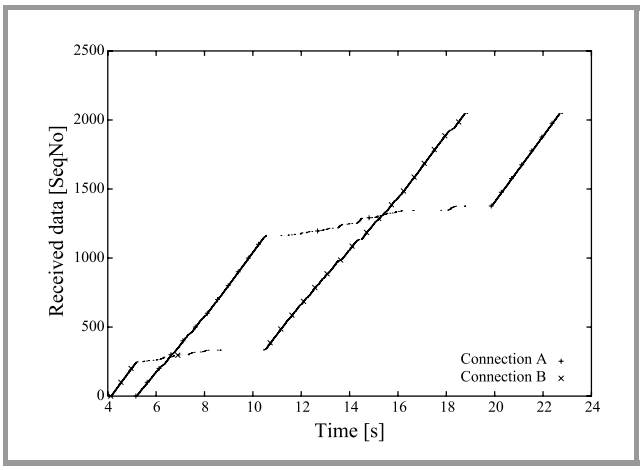


Fig. 10. Message Retraining model – controlled SNR.

In the following section, we employ two fairness indices to quantify the relative performance of each model against experimental data.

5. Comparison metrics

To make a quantitative comparison of the results obtained with each capture model, a fairness metric is required. Fairness in wireless networks can be a difficult quantity to define. In this context we require that each node is able to access the channel without sustained delay, and that no node

is able to monopolise the radio channel at the expense of other nodes. This should be independent of the physical network topology.

Following [19], we employ two fairness indices: Jain's fairness index, and a new index proposed in [19], the Kullback-Leibler fairness index. In each case, a sliding window method is used to calculate the fairness over a horizon of 200 frames. The window slides along the packet sequence indicating which node has successfully gained access to the channel, calculating an instantaneous value for each index. The average value is then calculated across the entire trace. The selection of this window size is motivated by the length of the traces. A window of 200 frames corresponds to 10% of the frames transferred over each connection.

As the trace records successfully acknowledged data, these results give an indication of the fairness associated with the data transfer at the transport layer, including effects from the MAC and PHY layers. We calculate fairness in this manner, as TCP is the most common transport protocol in use today, and any wireless PHY/MAC protocol must be expected to support competing TCP streams without imposing additional fairness characteristics.

5.1. Jain's fairness index

This index has been used widely in the literature to capture fairness characteristics in both congestion control [20] and wireless MAC protocols [19]. A perfectly fair distribution of channel access would result in a value of 1 for this index, though values above 0.95 are typically considered to indicate good fairness properties. The index F_j is defined in Eq. (4):

$$F_j = \frac{\left(\sum_{i=1}^{N} \rho_i\right)^2}{N \sum_{i=1}^{N} {\rho_i}^2}, \tag{4}$$

where ρ_i is the fractional share achieved by the ith connection, and N is the number of active connections. A value of 0.7 would imply that 30% of nodes were suffering significant unfairness.

5.2. Kullback-Leibler fairness index

The Kullback-Leibler fairness index was first proposed in [19]. The technique considers the distribution of channel access for each node as a probability distribution, $\tilde{\Gamma}$. The Kullback-Leibler distance $D\left(\Gamma||\tilde{\Gamma}\right)$, an entropy measure of the "distance" between two probability distributions, is calculated between the desired distribution Γ, and the measured distribution, $\tilde{\Gamma}$:

$$\begin{aligned} D\left(\Gamma||\tilde{\Gamma}\right) &= D\left(\left[\rho_1, \rho_2 \ldots \rho_n\right] || \left[\frac{1}{N}, \frac{1}{N} \cdots \frac{1}{N}\right]\right) = \\ &= \left(\sum_{i=1}^{N} \rho_i \log_2 \rho_i\right) + \log_2 N \end{aligned} \tag{5}$$

where: N is the number of nodes, and ρ_i the fractional share achieved by the ith node.

This measure provides an indication of the fairness in the system. A value of 0 corresponds to a perfectly fair system, with values below 0.05 typically indicating a system with good fairness properties.

Table 2
Fairness index comparison

Capture model	Experiment			
	stationary SNR		controlled SNR	
	Jain	K-L	Jain	K-L
None	0.80	0.23	0.80	0.23
Delay	0.80	0.29	0.73	0.39
Power	0.80	0.29	0.80	0.29
Hybrid	0.80	0.29	0.80	0.29
M-R	0.63	0.58	0.67	0.46
Trace data	0.68	0.52	0.62	0.68

Simulation of both the stationary and controlled signal power trials of Section 2 were performed and both fairness indices calculated. These results are then compared with the fairness indices from the measured traces. Table 2 illustrates the average fairness index for each capture model.

5.3. Discussion

The stationary signal power experiment illustrates that the delay, power and hybrid capture models provide an overestimate of the fairness observed experimentally. The Message Retraining model though slightly underestimating the fairness measured with both indices, provides an excellent indication of the fairness properties present in the experimental data.

The controlled signal power scenario represents a more challenging task for the capture models than the stationary signal power scenario. Varying signal conditions throughout the experiment are reflected in the measured trace, and should also be observed in a simulated trace. As illustrated in Section 4, this was not the case, and we would therefore expect the delay, power and hybrid models to significantly overestimate the fairness achieved by each connection. The results in Table 2 confirm this, with the delay, power and hybrid models all significantly overestimating the fairness present in the experimental data.

The Message Retraining model matches the experimental data according to Jain's index quite closely, though the Kullback-Leibler index is still shows significantly higher unfairness. This can be attributed to the long period in the experimental data (Fig. 3) between 15 and 32 seconds where no data is transferred. This is due to a significant TCP timeout on connection A. While the Message Retraining model resulted in a significant reduction in throughput during this time (Fig. 10), connection A was still able

to maintain sufficient throughput to prevent a long TCP timeout. This result also indicates the sensitivity of the Kullback-Leibler index.

6. Conclusions

Simulation plays an important role in performance evaluation of wireless MAC protocols. In this paper we have investigated the performance of a number of common modem capture models presented in literature, in terms of their ability to accurately reflect fairness properties of experimentally derived trace data. We have proposed a new capture model, which we show is able to model the dynamic fairness properties of the IEEE 802.11 PHY/MAC with varying signal conditions more accurately than the delay, power, or hybrid capture model.

Quantitative comparison between experimental trace data and simulation traces for each capture model using both Jain's fairness index and the Kullback-Leibler fairness index, illustrates than the delay, power, and hybrid capture models provide an overly optimistic estimate of the fairness afforded to the contending hidden connections. The Message Retraining model is shown to match the experimental data well.

Our results indicate that in cases where fairness is an important component of network performance, a more detailed capture model is required to reflect the impact of varying signal strength characteristics, and describe modem behaviour in a more complete manner.

References

[1] C. G. Ware, J. Judge, J. F. Chicharo, and E. Dutkiewicz, "Unfairness and capture behaviour in 802.11 ad hoc networks", in *Int. Conf. Commun.*, New Orleans, 2000, vol. 1.

[2] K. Cheun and S. Kim, "Joint delay-power capture in spread-spectrum packet radio networks", *IEEE Trans. Commun.*, vol. 46, no. 4, pp. 450–453, 1998.

[3] D. H. Davis and S. Gronemeyer, "Performance of slotted aloha random access with delay capture and randomised time of arival", *IEEE Trans. Commun.*, vol. COM-28, pp. 703–710, 1980.

[4] J. C. Arnbak, "Capacity of slotted aloha in Rayleigh fading channels", *JSAC*, vol. SAC-5, pp. 261–269, 1987.

[5] C. G. Ware, T. Wysocki, and J. Chicharo, "Hidden terminal jamming problems in IEEE 802.11 mobile ad hoc networks", in *Int. Conf. Commun.*, Helsinki, 2001.

[6] M. B. Pursley, "Performance evaluation for phase coded spread spectrum multiple access communication". Part 1. "System analysis", *IEEE Trans. Commun.*, vol. COM-25, no. 8, pp. 795–799, 1977.

[7] D. I. Kim, I. K. Kim, and R. A. Scholtz, "Counting collision free transmissions in common-code SSMA communications", *IEEE Trans. Commun.*, vol. 43, 1995.

[8] R. Mud, J. Boer, A. Kamerman, H. Van Driest, W. Diepenstraten, R. Kopmeiners, and H. Von Bokhorst, "Wireless LAN with enhanced capture provision". US patent, no. US5987033.

[9] M. Soroushnejad and E. Geraniotis, "Probability of capture and rejection of primary multiple access interference in spread spectrum networks", *IEEE Trans. Commun.*, vol. 39, no. 6, pp. 986–994, 1991.

[10] T. Nandagopal, T.-E. Kim, X. Gao, and V. Bharghavan, "Achieving MAC layer fairness in wireless packet networks", in *ACM Mobicom*, Boston, MA, 2000.

[11] B. Bensaou, Y. Wang, and C. C. Ko, "Fair medium access in 802.11 based wireless ad hoc networks", in *Mob. Netw. Comput. (Mobihoc. 2000)*, Boston, MA, 2000, pp. 99–106.

[12] M. Gerla, K. Tang, and R. Bagrodia, "TCP performance in wireless multihop networks", in *2nd IEEE Workshop Mob. Comput. Syst. Appl.*, 1999, vol. 1, pp. 41–50.

[13] M. Gerla, R. Bagrodia, K. Tang, and L. Wang, "TCP over wireless multihop protocols: simulation and experiments", in *IEEE Int. Conf. Commun.*, Vancouver, Canada, 1999, vol. 1.

[14] K. Tang and M. Gerla, "Fair sharing of MAC under TCP in wireless ad hoc networks", in *Multimed. Mob. Teletraff. Wirel. Net. (MMT'99)*, Venice, Italy, 1999.

[15] Institution of Electrical and Electronic Engineers. Part 11: "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications", 1997.

[16] Institution of Electrical and Electronic Engineers. Part 11: "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications, higher speed physical layer extension in the 2.4 GHz band", 1999.

[17] Institution of Electrical and Electronic Engineers. Part 11: "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications, higher speed physical layer in the 5 GHz band", 1999.

[18] *tcpdump*, Lawrence Berkeley Laboratories, 1999.

[19] C. E. Koksal, H. Kassab, and H. Balakrishnan, "An analysis of short term fairness in wireless media access protocols: extended version of short paper", in *ACM Sigmetr.*, 2000.

[20] R. Jain, D. Chiu, and W. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems". Tech. Rep. TR-301, Dec. 1984.

Christopher Ware received the B.Sc. degree in Physics, and the B.E. degree (Hons I) in Electrical Engineering from the University of Wollongong, Australia in 1998. From 1997 to 1998, he was a Telecommunications Engineer with Alcatel Australia. In 1998 he joined the Telecommunications and Information Technology Research Institute at the University of Wollongong where he is currently a Research Assistant. He is also a Ph.D. candidate in the same department. His current research interests include wireless MAC protocol design, propagation and error models, and QoS in wireless networks.

University of Wollongong
School of Electrical, Computer
and Telecommunications Engineering
Northfields Av.
Wollongong NSW 2522, Australia

Joe Chicharo received his B.E. degree (Hons 1) and Ph.D. in Electrical Engineering from the University of Wollongong, Australia, in 1983 and 1990 respectively. He is a Fellow of the Institution of Engineers Australia and a Senior Member of the IEEE. Prior to joining the University of Wollongong Joe was employed (1977 to 1982) as a Degree Trainee by BHP Sheet and Coil Products Steel Division at Port Kembla, Australia. From 1983 to 1985 he was employed as an Electrical Engineer and then as Senior Engineer by the same company where he gained considerable industrial experience primarily

in automation and systems engineering as well as large capital project management. Joe has served as Managing Director of the Telecommunications and Information Technology Research (TITR) Institute, University of Wollongong since 1999. He is also the coordinator of the Switched Networks Research Centre (SNRC). SNRC has been a Telstra funded centre of excellence since 1990. Joe has been appointed as Research Director of the CRC for Smart Internet Technology. His current academic interests include a variety of research issues associated with the Internet including access technologies, small office/home office Internet enabled services, ad hoc and location aware networking, network traffic modelling and control. His other major research activity has been in the area of digital signal processing dealing with adaptive systems and speech processing issues. Joe has successfully supervised many post-graduate students at both Ph.D. and Honours Masters levels in the areas of Telecommunications and Digital Signal Processing. He has served as an Associate Editor of the IEEE Transactions on Circuits and Systems, II.

University of Wollongong
School of Electrical, Computer
and Telecommunications Engineering
Northfields Av.
Wollongong NSW 2522, Australia

Tadeusz A. Wysocki – for biography, see this issue, p. 27.

Paper

Error statistics for concatenated systems on non-renewal time-varying channels

Cecilio Pimentel

Abstract — The statistics of the error process generated by a discrete super channel formed by the concatenation of a constrained encoder, a non-renewal finite state channel (FSC), and a constrained decoder is studied in this paper. First, recursions are developed for the error weight distribution. This statistics is relevant to the design of coding schemes and interleaving in concatenated systems. We also study the renewal nature of the residual error process as modified by the constrained decoder. Ferreira *et al.* conjectured that if the channel model is a renewal FSC, the super channel can be modeled as a similar renewal model. We use a statistics called the multigap distribution to analytically disprove this hypothesis. Furthermore, the effect of interleaving is investigated from a new perspective using the variance of the multigap distribution.

Keywords — finite state channels, multigap distribution, constrained codes, concatenated systems, Gilbert-Elliott channels, combinatorial methods.

1. Introduction

Finite state channels are mathematical models based on probabilistic or deterministic function of Markov chains that are capable of modeling the correlated error sequence produced by a broad class of communication channels with memory. Thus, for example, in a binary system the error sequence is the pairwise modulo two sum of the input $\{x_k\}_{k=1}^{\infty}$ and output $\{y_k\}_{k=1}^{\infty}$ sequence of the channel defined as follows. At the kth time interval, the error bit e_k is equal to zero (indicating no error) if $x_k = y_k$, or e_k is equal to one (indicating an error) if $x_k \neq y_k$.

In several applications, such as magnetic and optical storage systems, the communication system uses the concatenation of forward error correcting codes and constrained codes. The constraints introduced into the transmitted sequence offer the possibility of achieving a desired spectral shaping, reducing the intersymbol interference and improving the synchronization ability of communication systems. The process of selecting an appropriate error control coding scheme should take into account the error statistics of the super channel in order to correct or detect a specific set of most probable error sequences produced by the super channel.

Ferreira *et al.* [1] introduced the problem of calculating a statistics known as error-free run distribution for a super channel comprised of a certain class of constrained codes and a burst channel modeled according to a renewal Fritchman channel with one error state. In such renewal channels, sequences of zeros (also known as gap intervals) before and after an error are statistically independent and identically distributed random variables. The renewal assumption simplifies the analysis of the model, but it can be demonstrated that many communication systems present some dependence in the occurrence of successive gap intervals [2, 3]. The Gilbert-Elliott channel and Fritchman channels with more than one error state are, for example, non-renewal FSC models. The characterization of the residual error sequence at the output of the super channel with embedded non-renewal FSC models is of interest in this paper.

The development of mathematical tools to use with general non-renewal FSC models for system performance evaluation has been considered in the literature [4 – 6]. In this paper, we extend the general framework proposed in [6] to compute the statistics of the error sequence at the output of the super channel as a function of the inner model parameters. The main idea is to find a fractional generating series whose coefficients of its series expansion yields the statistics of interest. This approach allows the development of generating series and recursions for certain statistics for which no previous methods existed.

This paper is organized as follows. Section 2 contains a brief review of FSC models. The characterization of the residual error sequence at the output of the super channel is the main subject of Section 3. First, we derive recursion for a statistics called error weight distribution, that is, the probability of m errors in a block of length n. This statistics is relevant to the design of coding schemes and interleaving in concatenated systems. Next, we study the multigap distribution of the super channel. This statistics has been used as a test of non-renewalness of the error process and is capable of revealing interesting properties of burst channels, such as, the gap length spread, and the correlation coefficient between gap intervals. It is analytically proved in this section that the pair constrained encoder/decoder may convert a renewal burst channel into a non-renewal super channel. Conclusions are summarized in Section 4.

We adopt the following notation throughout this work. The matrices $\mathbf{I}$ and $\mathbf{1}$ stand for the identity matrix and a column vector of ones. $\mathbf{E}\{E_k\}$ stands for the expected value of the random variable E_k. If s and z are commutative indeterminates, $[s^k z^n]\,P(s,z)$ denotes the coefficient of $s^k z^n$ in the formal power series $P(s,z)$.

2. Model description

Consider $\{S_k\}_{k=0}^{\infty}$ a stationary Markov chain with state space $\mathcal{N}_N = \{0, 1, \ldots, N-1\}$, transition probability matrix $\mathbf{P}$, and stationary probability row vector $\boldsymbol{\pi}$. At the kth time interval, the chain makes a transition from state $S_{k-1}=i$ to $S_k = j$ with probability $p_{i,j}$ and generates an error bit e_k (independent of i), with probability $P(E_k = e_k \mid S_k = j)$. Define two $N \times N$ matrices, $\mathbf{P}(e_k)$, $e_k \in \{0,1\}$, whose (i,j)th entry is $P(E_k = e_k \mid S_k = j)P(S_k = j \mid S_{k-1} = i)$, which is the probability that the output symbol is e_k when the chain makes a transition from state i to j. The probability of an error sequence of length n, may be expressed in a matrix form as:

$$P(e_1 e_2 \cdots e_n) = \boldsymbol{\pi} \left(\prod_{k=1}^{n} \mathbf{P}(e_k) \right) \mathbf{1}.$$

The matrices $\mathbf{P}(0)$, $\mathbf{P}(1)$ and $\boldsymbol{\pi}$ for the Gilbert-Elliott channel (GEC) [7] illustrated in Fig. 1 are given below:

$$\mathbf{P}(0) = \begin{bmatrix} (1-Q)(1-g) & Q(1-b) \\ q(1-g) & (1-q)(1-b) \end{bmatrix}; \quad (1)$$

$$\mathbf{P}(1) = \begin{bmatrix} (1-Q)g & Qb \\ qg & (1-q)b \end{bmatrix}; \quad (2)$$

$$\boldsymbol{\pi} = [\pi_0 \, \pi_1] = \left[\frac{q}{q+Q} \; \frac{Q}{q+Q} \right]. \quad (3)$$

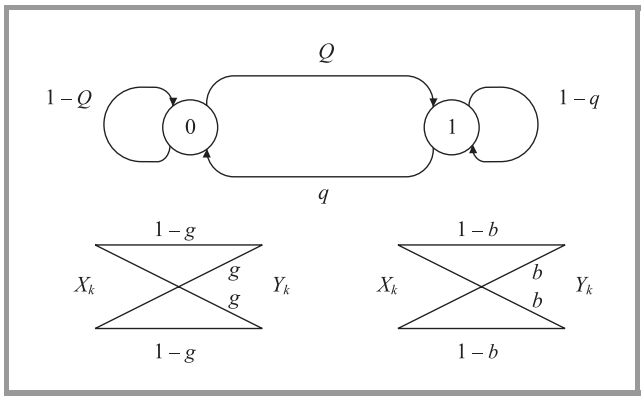


Fig. 1. Gilbert-Elliott model for burst channels.

An FSC model can also be described as a deterministic function of a Markov chain. In the Fritchman channel [8], the state space $\mathcal{N}_N$ is partitioned into two disjoint subsets, $\mathcal{A}_0 = \{0, \ldots, N-2\}$ (the good states), and $\mathcal{A}_1 = \{N-1\}$ (the bad state). The error bit, e_k, is a *deterministic* function of the current state S_k, and assumes the value $e_k = 0$ (no error) if $S_k \in \mathcal{A}_0$ or $e_k = 1$ (error) if $S_k \in \mathcal{A}_1$. A model with K good states and 1 bad state is denoted by $(K,1)$-FC. In particular, the matrices $\mathbf{P} = \mathbf{P}(0) + \mathbf{P}(1)$, $\mathbf{P}(0)$, $\mathbf{P}(1)$, for the $(2,1)$-FC model are given by:

$$\mathbf{P} = \left[\begin{array}{cc|c} \lambda_0 & 0 & 1-\lambda_0 \\ 0 & \lambda_1 & 1-\lambda_1 \\ \hline p_{2,0} & p_{2,1} & 1-p_{2,0}-p_{2,1} \end{array} \right]; \quad (4)$$

$$\mathbf{P}(0) = \left[\begin{array}{cc|c} \lambda_0 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ \hline p_{2,0} & p_{2,1} & 0 \end{array} \right]; \quad (5)$$

$$\mathbf{P}(1) = \left[\begin{array}{cc|c} 0 & 0 & 1-\lambda_0 \\ 0 & 0 & 1-\lambda_1 \\ \hline 0 & 0 & 1-p_{2,0}-p_{2,1} \end{array} \right]. \quad (6)$$

Tsai modeled a fading HF communication channel using the following (2,1)-FC model [8]:

$$\mathbf{P} = \left[\begin{array}{cc|c} 0.99911 & 0 & 0.00089 \\ 0 & 0.73644 & 0.26356 \\ \hline 0.36258 & 0.58510 & 0.05232 \end{array} \right]. \quad (7)$$

3. Error statistics

Two statistics of the residual error sequence produced by the super channel are derived in this section. The error weight distribution and the multigap distribution.

3.1. Error weight distribution

The error weight distribution is the probability of an FSC generates exactly m consecutive errors in a block of length n. This probability is denoted by $P(m,n)$. An expression for $P(m,n)$ can be obtained by first finding the following bivariate generating series:

$$H_P(s,z) = \sum_{n=0}^{\infty} \sum_{m=0}^{n} P(m,n) s^m z^n, \quad (8)$$

where the indeterminates s and z mark the number of ones and the length of the error sequence, respectively. Our interest in this series relies on the fact that the coefficient of $s^m z^n$ is the quantity of interest, that is, $P(m,n) = [s^m z^n] H_P(s,z)$. The generating series $H_P(s,z)$ for the error process $\{e_k\}_{k=1}^{\infty}$ of FSC models is given in [6] as:

$$H_P(s,z) = \boldsymbol{\pi}(\mathbf{I} - \mathbf{P}(1)sz - \mathbf{P}(0)z)^{-1}\mathbf{1}. \quad (9)$$

Upon substitution of the matrices (1)-(3) into (9), we express $H_P(s,z)$ for the GEC as the ratio of two polynomials in s and z:

$$H_P(s,z) = \frac{1 + c_{1p}z + c_{2p}zs}{1 + c_1 z + c_2 zs + c_3 z^2 + c_4 z^2 s + c_5 z^2 s^2}, \quad (10)$$

where

$$\begin{aligned} c_{1p} &= [qQ(2-b-g) - q(1-q)(1-b) + \\ &\quad -Q(1-Q)(1-g)]/(q+Q); \\ c_{2p} &= [qQ(b+g) - Q(1-Q)g - q(1-q)b]/(q+Q), \end{aligned}$$

and

$$\begin{aligned} c_1 &= (Q(1-g)+q(1-b)-(2-g-b));\\ c_2 &= -(b(1-q)+g(1-Q));\\ c_3 &= (1-b)(1-g)(1-q-Q);\\ c_4 &= (1-q-Q)(b+g-2gb);\\ c_5 &= (1-q-Q)gb. \end{aligned} \tag{11}$$

It is simple go from a generating series to recurrence formulas, which provides a rapid computational scheme for the problem. The denominator polynomial of $H_P(s,z)$ is responsible for the recurrence relation, while the numerator polynomial defines the initial conditions. From the generating series (10) we derive the following 6-term recurrence formula for $P(m,n)$ for the GEC:

$$\begin{aligned} P(m,n) = &-c_1P(m,n-1)-c_2P(m-1,n-1)+\\ &-c_3P(m,n-2)-c_4P(m-1,n-2)+\\ &-c_5P(m-2,n-2), \end{aligned} \tag{12}$$

for $m \geq 0$, $n > 1$, where the coefficients $\{c_i\}_{i=1}^5$ are given by (11). The initial conditions are:

$$\begin{aligned} P(m,n) &= 0, \text{ for } m,n<0,\ m>n;\ \ P(0,0)=1;\\ P(0,1) &= \boldsymbol{\pi}\mathbf{P}(0)\mathbf{1} = c_{1p}-c_1;\\ P(1,1) &= \boldsymbol{\pi}\mathbf{P}(1)\mathbf{1} = c_{2p}-c_2. \end{aligned}$$

We now turn to the calculation of $P(m,n)$ for the error sequence produced by the super channel. This statistics is sensitive to the operation of the constrained decoder. The next examples consider specific constrained codes.

Example 1 (1/2-rate NRZ Miller code). The encoder and decoder operations of the 1/2-rate nonreturn-to-zero (NRZ) Miller code are explained in [1]. At the kth interval, the bit c_k is the input to the constrained encoder yielding two output bits, $x_{k,1}x_{k,2}$, which are the inputs to the FSC model. The two output bits of the channel, $y_{k,1} = x_{k,1} \oplus e_{k,1}$ and $y_{k,2} = x_{k,2} \oplus e_{k,2}$, where $\oplus$ is modulo 2 addition, are the input to the constrained decoder. The decoder maps these two bits onto the symbol r_k, according to the following rule:

$$\begin{aligned} r_k = y_{k,1} \oplus y_{k,2} &= (x_{k,1} \oplus x_{k,2}) \oplus (e_{k,1} \oplus e_{k,2});\\ &= c_k \oplus z_k, \end{aligned}$$

where $z_k = e_{k,1} \oplus e_{k,2}$ is the the binary error sequence for the super channel. It is clear that z_k is equal to 0, that is $r_k = c_k$, if $e_{k,1}e_{k,2}$= 00 or 11. Otherwise, z_k is equal to 1, that is $r_k \neq c_k$, if $e_{k,1}e_{k,2}$ = 01 or 10. From this explanation, we conclude that the probability $P(m,n)$ for the error sequence $\{z_k\}_{k=1}^{\infty}$ for the 1/2-rate NRZ Miller code is expressed in a form similar to (9):

$$P(m,n) = [s^m z^n]\,\boldsymbol{\pi}(\mathbf{I}-\mathbf{P}'(1)sz-\mathbf{P}'(0)z)^{-1}\mathbf{1}, \tag{13}$$

where the matrices $\mathbf{P}'(0)$ and $\mathbf{P}'(1)$ are:

$$\begin{aligned} \mathbf{P}'(0) &= \mathbf{P}(0)\,\mathbf{P}(0)+\mathbf{P}(1)\,\mathbf{P}(1);\\ \mathbf{P}'(1) &= \mathbf{P}(0)\,\mathbf{P}(1)+\mathbf{P}(1)\,\mathbf{P}(0). \end{aligned} \tag{14}$$

Thus

$$\begin{aligned} P(m,n) = [s^m z^n]\boldsymbol{\pi}\{\ &\mathbf{I}-[\,\mathbf{P}(0)\,\mathbf{P}(1)+\mathbf{P}(1)\,\mathbf{P}(0)]sz+\\ &-[\,\mathbf{P}(0)\,\mathbf{P}(0)+\mathbf{P}(1)\mathbf{P}(1)]z\}^{-1}\mathbf{1}, \end{aligned} \tag{15}$$

yielding the same recursion formula given by (12), but with the following coefficients:

$$\begin{aligned} c_1 &= -(1-q)^2(1-2b+2b^2)-(1-Q)^2(1-2g+2g^2)+\\ &\quad -2qQ(1-b-g+2gb);\\ c_2 &= -2(1-Q)^2(1-g)g-2qQ(b+g-2gb)+\\ &\quad -2(1-q)^2(1-b)b;\\ c_3 &= (1-Q)^2(1-q)^2(1-2b+2b^2)(1-2g+2g^2)+\\ &\quad -qQ(1-q)(1-Q)[b^2(1-g)^2+2g^2b^2+\\ &\quad +2(1-g)^2(1-b)^2+2(1-g)(1-b)gb+\\ &\quad +g^2(1-b)^2]+q^2Q^2(1-b-g+2gb)^2;\\ c_4 &= 2[b(1-b)(1-2g+2g^2)+\\ &\quad +g(1-g)(1-2b+2b^2)](1-q-Q)^2;\\ c_5 &= 4(1-Q)^2(1-g)g(1-q)^2(1-b)b+\\ &\quad +q^2Q^2(b+g-2gb)^2+\\ &\quad -(1-Q)Q(1-q)q[6(1-g)(1-b)bg+\\ &\quad +(1-g)^2b^2+g^2(1-b)^2], \end{aligned}$$

and initial conditions:

$$\begin{aligned} P(m,n) &= 0, \text{ for } m,n<0,\ m>n;\ \ P(0,0)=1;\\ P(0,1) &= \boldsymbol{\pi}\mathbf{P}'(0)\mathbf{1};\\ P(1,1) &= \boldsymbol{\pi}\mathbf{P}'(1)\mathbf{1}, \end{aligned}$$

where the matrices $\mathbf{P}'(0)$ and $\mathbf{P}'(1)$ are given by (14).

Example 2 (Systematic 1/2-rate Miller code). We consider now a systematic 1/2-rate Miller code [1]. At each interval, the constrained decoder receives a pair of bits, $y_{k,1}y_{k,2}$, and the decoded bit r_k is read directly from the second bit $y_{k,2}$. The first received bit $y_{k,1}$ is discarded in the decoding process. The error bit for the super channel z_k is equal to one if $e_{k,1}e_{k,2} = (0 \cup 1)1$, or z_k is equal to zero if $e_{k,1}e_{k,2} = = (0 \cup 1)0$. $P(m,n)$ is given by formula (13), where

$$\begin{aligned} \mathbf{P}'(0) &= \mathbf{P}\mathbf{P}(0);\\ \mathbf{P}'(1) &= \mathbf{P}\mathbf{P}(1). \end{aligned} \tag{16}$$

Thus

$$P(m,n) = [s^m z^n]\,\boldsymbol{\pi}(\mathbf{I}-\mathbf{P}\mathbf{P}(1)sz-\mathbf{P}\mathbf{P}(0)z)^{-1}\mathbf{1}. \tag{17}$$

The recursion formula for $P(m,n)$ for the GEC is given by (12) with the following coefficients:

$$\begin{aligned} a_1 &= -(1-q)^2(1-b)-qQ(2-b-g)+\\ &\quad -(1-Q)^2(1-g);\\ a_2 &= (2-q-Q)(Qg+qb)-(b+g);\\ a_3 &= -2(1-Q)Q(1-g)(1-q)q(1-b)+\\ &\quad +(1-Q)^2(1-g)(1-q)^2(1-b)+\\ &\quad +Q^2q^2(1-b)(1-g);\\ a_4 &= (b+g-2gb)[-2Q(1-Q)q(1-q)+\\ &\quad +(1-Q)^2(1-q)^2+Q^2q^2];\\ a_5 &= bg[(1-q)^2-Q(2-2q-Q)], \end{aligned} \tag{18}$$

and initial conditions:

$$\begin{aligned} P(m,n) &= 0, \text{ for } m,n<0,\ m>n;\ \ P(0,0)=1;\\ P(0,1) &= \boldsymbol{\pi}\mathbf{P}'(0)\mathbf{1};\\ P(1,1) &= \boldsymbol{\pi}\mathbf{P}'(1)\mathbf{1}, \end{aligned} \tag{19}$$

where the matrices $\mathbf{P}'(0)$ and $\mathbf{P}'(1)$ are given by Eqs. (16). To exemplify an application of recursion given by formulae (12), (18) and (19) we consider a situation where an error correcting (15,7) BCH code with error correction capability $t=2$ is used to correct the errors produced by a super channel. Figure 2 shows the probability of codeword error, PCE = $\sum_{m=t+1}^{n} P(m,n)$, as a function of the average burst length of the GEC, $\lambda = 1/q$. PCE for a GEC alone is also shown in the figure. The following channel parameters are held fixed $q/Q = 20$, $b = 0.3$, $g = 1\cdot 10^{-3}$.

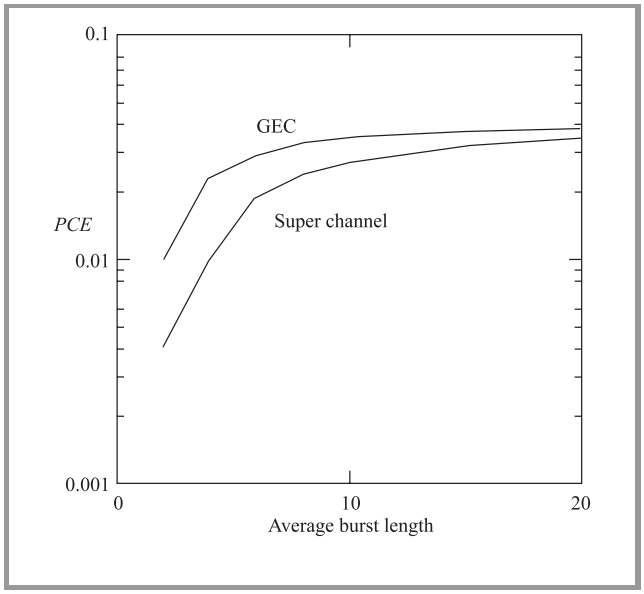


Fig. 2. Probability of codeword error as a function of the average burst length for a (15,7) BCH code on a super channel. The super channel is the concatenation of the systematic 1/2-rate Miller code and the GEC model.

3.2. The multigap distribution

It was conjectured in [1] that the error sequence $\{z_k\}_{k=1}^{\infty}$ generated by a super channel comprised of an 1/2-rate NRZ Miller code and a renewal Fritchman channel can be modeled as a renewal Fritchman channel. In this section, we develop a statistics called the variance of the multigap distribution to disprove this hypothesis analytically, by showing the non-renewal nature of this super channel.

The length of a gap is the number of zeros between two errors plus one (the last error is included). The error process $\{E_k\}_{k=1}^{\infty}$ can be regarded as a sequence of gaps $\{G_k\}_{k=1}^{\infty}$, where G_k is the length of the kth gap. For example, the error sequence of length 9, $\mathbf{E}_9 = 110000001$ corresponds to the gap sequence $\mathbf{G}_9 = 117$ (by assumption, $E_0 = 1$). Let the random variable $G^r = \sum_{i=k}^{k+r-1} G_k$ be the sum of r consecutive gap lengths. The multigap length distribution, denoted as $M(r,l)$, is defined as $M(r,l) = P(G^r = l)$.

If the error process is renewal, this means that $\{G_k\}_{k=1}^{\infty}$ are independent random variables, then the variance of G^r is $Var(G^r) = r\,Var(G_1)$. Therefore, a linear growth of $Var(G^r)$ with r is sufficient to show the renewal nature of the channel. The problem of finding an expression for the multigap distribution for FSC models was considered in [6]. It is shown in [6] that the generating series $H_M(s,z)$ for $M(r,l)$ is:

$$\begin{aligned} H_M(s,z) &= \sum_{r,l} M(r,l)\, s^r z^l;\\ &= \tfrac{1}{P(1)}\,\boldsymbol{\pi}\mathbf{P}(1)(\mathbf{I}-(\mathbf{I}-\mathbf{P}(0)z)^{-1}\mathbf{P}(1)sz)^{-1}\mathbf{1}. \end{aligned} \tag{20}$$

The variance of G^r, denoted as $Var(G^r)$, can be expressed as:

$$Var(G^r) = [s^r]\left\{\frac{\partial^2 H_M(s,z)}{\partial z^2}\right\}_{z=1} + \mathbf{E}\{G^r\}(1-\mathbf{E}\{G^r\}). \tag{21}$$

To study the renewal property of communication channels, Adoul [9] defined a quantity called variation coefficient $K(r)$ as:

$$Var(G^r) = K(r)\, Var_{BSC}(G^r), \tag{22}$$

where $Var_{BSC}(G^r) = r(1-P(1))/P(1)^2$ is the variance of G^r for the memoryless binary symmetric channel (BSC) channel with crossover probability $P(1)$. It is important to notice that for renewal processes $K(r) = K(1)$, for all r, that is, the curve $K(r)$ versus r is a constant for all r. An expression for $K(r)$ for the renewal (2,1)-FC model is given below:

$$\begin{aligned} K(r) &= (1+\lambda_0)(1+\lambda_1)[p_{2,1}(1-\lambda_0)+(1-\lambda_1)p_{2,0}]+\\ &-[(1-\lambda_0)p_{2,1}+(1-\lambda_1)p_{2,0}]^2/\,[(1-\lambda_0)(1-\lambda_1)+\\ &+p_{2,0}(1-\lambda_1)+p_{2,1}(1-\lambda_0)][p_{2,0}(1-\lambda_1)+p_{2,1}(1-\lambda_0)]. \end{aligned}$$

We now investigate the renewal property of the error process $\{z_k\}_{k=1}^{\infty}$ of the super channel. To address this problem for a specific constrained code, we calculate the following generating series:

$$H_V(s) = \sum_{r=0}^{\infty} Var(G^r)\, s^r, \tag{23}$$

where $Var(G^r)$ is calculated for the process $\{z_k\}_{k=1}^{\infty}$ using Eqs. (20) and (21) with the matrices $\mathbf{P}(0)$ and $\mathbf{P}(1)$ replaced by $\mathbf{P}'(0)$ and $\mathbf{P}'(1)$ given by Eqs. (14) and (16). If the process is renewal, we have:

$$H_V(s) = Var(G^1)\frac{s}{(1-s)^2} = Var(G^1)\sum_{r=0}^{\infty} r\,s^r. \tag{24}$$

We consider first the super channel comprised of an 1/2-rate NRZ Miller code and a (2,1)-FC with matrices $\mathbf{P}(0)$ and

$\mathbf{P}(1)$ given by Eqs. (5) and (6). Using the matrices $\mathbf{P}'(0)$ and $\mathbf{P}'(1)$ given by Eqs. (14) we found that $H_V(s)$ for this super channel is of the form:

$$H_V(s) = \frac{b_1 s}{(1-s)^2}\frac{(b_2+b_3 s)}{(b_4+b_5 s)}, \tag{25}$$

where $\{b_i\}_{i=1}^{5}$ are distinct constants depending on the FSC parameters. Therefore, we proved that the super channel under consideration is non-renewal. In [1] a renewal (2,1)-FC model was proposed to represent this super channel when the inner channel is the Tsai's (2,1)-EFC model given by Eq. (7). Numerical values for $K(r)\cdot r$ for this super channel is shown in Table 1. The values of $K(r)$ in the table differ only in the third significant digit, indicating that the renewal approximation is valid for these specific parameters.

Table 1
Variation coefficient $K(r)$ versus the number of consecutive gap lengths r. The super channel is the concatenation of the 1/2-rate NRZ Miller code and the Tsai's (2,1)-EFC model given by Eq. (7)

r	$K(r)$
1	4.21862
2	4.21791
3	4.21764
4	4.21750
10	4.21728
20	4.21718
30	4.21715

When we consider the systematic 1/2-rate Miller code with matrices $\mathbf{P}'(0)$ and $\mathbf{P}'(1)$ given by Eqs. (16) the generating series $H_V(s)$ satisfies Eq. (24) and this super channel is renewal. Noticed that, in this case, the process $\{z_k\}_{k=1}^{\infty}$ is a sample of the error sequence $\{e_k\}_{k=1}^{\infty}$, or $z_k = e_{2k}$. In general, we can show that if the samples are spaced l intervals apart, or $z_k = e_{lk}$, the renewal property is maintained. Table 2 shows that values of $K(1)$ versus l for the sampled process $z_k = e_{lk}$, when the inner channel is the (2,1)-FC model given by Eq. (7).

Adoul [9] defined a process whose $K(1)$ is greater than one as a *more variable* process, in the sense that the gap lengths spread widely from their mean value (errors have a trend to be clustered). The further $K(1)$ is from 1, the more pronounced is this trend. Table 2 shows that $K(1) = 4.443$ for the (2,1)-FC model ($l = 1$), and $K(1)$ decreases to 2.761 for the super channel with the systematic 1/2-rate NRZ Miller code ($l = 2$). As the value of l increases, the process $\{z_k\}_{k=1}^{\infty}$ tends to become memoryless and $K(1)$ tends to 1. If we encompass the FSC model with an interleaving and a deinterleaving with finite interleaving depth l, we can regard the sampled process $z_k = e_{lk}$ as the error sequence at each row of the deinterleaver. We can use the results in Table 2 to investigate the ideal value of l that renders the channel memoryless. Numerical values for $K(r)\cdot r$

Table 2
$K(1)$ versus interleaving depth l. The inner channel is the (2,1)-FC model given by Eq. (7)

l	$K(1)$
1	4.443
2	2.761
3	2.100
4	1.790
5	1.600
10	1.235
20	1.068
30	1.025
40	1.0099

for the GEC and the concatenation of the 1/2-rate NRZ Miller code and the GEC are shown in the solid curves of Fig. 3. The channel parameter considered are $Q = 4\cdot 10^{-6}$, $q = 4.7\cdot 10^{-4}$, $b = 0.3$ and $g = 1\cdot 10^{-3}$. The positive increment $K(2) - K(1)$ indicates that both channels have positive correlation between gap intervals. The analysis of the super channel may be simplified if we define a new FSC model that characterizes the error structure of the super channel. The triangle symbols in the figure shows the behavior of the $K(r)$ for a GEC that represents the super channel under consideration, where we found that the parameters of the new GEC are $Q = 8\cdot 10^{-6}$, $q = 9.3\cdot 10^{-4}$, $b = 0.42$ and $g = 2\cdot 10^{-3}$. These results differed from the solid curves in the third significant digit.

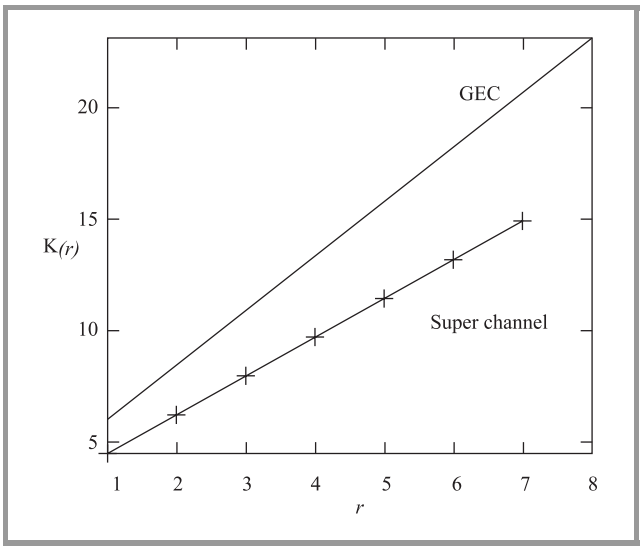


Fig. 3. Variation coefficient $K(r)$ as a function of the number of consecutive gap lengths r. The super channel is the concatenation of the 1/2-rate NRZ Miller code and the GEC model.

Equally accurate results were obtained for the $P(m,n)$ statistics. From this analysis, we conclude that the GEC represents the super channel with high accuracy.

4. Conclusions

We have studied the statistics of a binary error process generated by a super channel comprised of a specific constrained code and a general non-renewal FSC model. The theory proposed in [6] is extended to compute the generating series for the error weight distribution in terms of the channel matrices $\boldsymbol{\pi}$, $\mathbf{P}$, $\mathbf{P}(0)$ and $\mathbf{P}(1)$, for two specific constrained codes. The variance of the multigap distribution was employed to prove that the incorporation of a constrained code into the system may convert a renewal burst channel into a non-renewal super channel. This analysis was applied to a sampled error process in order to investigate the ideal value of the interleaving depth. Finally, our results suggest that an FSC could be directly used to model a super channel, simplifying the analysis of the overall system.

Acknowledgements

The author acknowledges partial support of this research by the Brazilian National Council for Scientific and Technological Development (CNPq) under grant 300987/96-0. The author would like to thank the anonymous reviewers for providing valuable comments on an earlier version of the manuscript.

References

[1] H. C. Ferreira, C. S. Coetzee, and M. A. Herro, "Mathematical models for super channels with imbedded constrained codes", *IEEE Trans. Inform. Theory*, vol. 39, pp. 1094–1100, 1993.

[2] L. N. Kanal and A. R. K. Sastry, "Models for channels with memory and their applications to error control", *Proc. IEEE*, vol. 66, pp. 724–744, 1987.

[3] W. Turin and R. M. Mohan Sondhi, "Modeling error sources in digital channels", *IEEE J. Select. Areas Commun.*, vol. 11, pp. 340–347, 1993.

[4] J. R. Yee and E. J. Weldon Jr., "Evaluation of the performance of error-correcting codes on a Gilbert channel", *IEEE Trans. Commun.*, vol. 43, pp. 2316–2323, 1995.

[5] M. Zorzi and R. R. Rao, "On the statistics of block errors in burst channels", *IEEE Trans. Commun.*, vol. 45, pp. 660–667, 1997.

[6] C. Pimentel and Ian F. Blake, "Enumeration of Markov chains and burst error statistics for finite state models", *IEEE Trans. Veh. Technol.*, vol. 48, no. 2, pp. 415–428, 1999.

[7] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels", *Bell Syst. Techn. J.*, vol. 42, pp. 1977–1997, 1963.

[8] B. D. Fritchman, "A binary channel characterization using partitioned Markov chains", *IEEE Trans. Inform. Theory*, vol. 13, pp. 221–227, 1967.

[9] J. A. Adoul, "Error intervals and cluster density in channel modeling", *IEEE Trans. Inform. Theory*, vol. 20, pp. 125–129, 1974.

Cecilio Pimentel was born in Recife, Brazil, in 1966. He received the B.S.E.E. degree from the Federal University of Pernambuco, Brazil, in 1987; the M.S.E.E. degree from the Catholics University of Rio de Janeiro, Brazil, in 1990; and the Ph.D. degree in electrical engineering from the University of Waterloo, Ontario, Canada, in 1996. Since October 1996, he has been with the Department of Electronics and Systems at the Federal University of Pernambuco, where he is currently an Associate Professor. His research interests include digital communications, information theory, and error correcting coding.
e-mail: cecilio@npd.ufpe.br
Department of Electronics and Systems
Federal University of Pernambuco
P.O. Box 7800, 50711-970 Recife - PE - Brazil

Paper

Stop criteria for retransmission termination in soft-combining algorithms

Hans-Jürgen Zepernick and Manora Caldera

Abstract — Soft-combining algorithms use retransmissions of the same codeword to improve the reliability of communication over very noisy channels. In this paper, soft-outputs from a maximum *a posteriori* (MAP) decoder are used as *a priori* information for decoding of retransmitted codewords. As all received words may not need the same number of retransmissions to achieve satisfactory reliability, a stop criterion to terminate retransmissions needs to be identified. As a first and very simple stop criterion, we propose an algorithm which uses the sign of the soft-output at the MAP decoder. The performance obtained with this stop criterion is compared with the one assuming a genius observer, which identifies otherwise undetectable errors. Since this technique needs always a particular number of initial retransmissions, we exploit cross-entropy between subsequent retransmissions as a more advanced but still simple stop criterion. Simulation results show that significant performance improvement can be gained with soft-combining techniques compared to simple hard or soft decision decoding. It also shows that the examined stop criteria perform very close to the optimistic case of a genius observer.

Keywords — MAP decoder, soft-combining, retransmission termination, cross-entropy.

1. Introduction

In order to provide efficient and reliable data transmission over a noisy communication channel, error control coding techniques are employed as an essential part in almost every modern digital communication system and in particular in mobile radio systems. In addition, hybrid automatic repeat request (ARQ) schemes are used to further improve reliability in a noisy channel. A hybrid ARQ scheme basically uses an error-correcting code to detect and if necessary to correct transmission errors. If the error pattern is detectable but not correctable, the receiver discards the received word and asks for a retransmission of that particular codeword. Unfortunately in this case, the whole effort put into the decoding process and the information gained from that failed decoding attempt is completely lost. This can be avoided by combining several repeated codewords at the output of a noisy channel, for example, using a maximum likelihood decoder [1]. Soft-combining algorithms incorporate reliability information into the decoding process and use soft values on a symbol-by-symbol basis. In this case, it is beneficial to exploit soft-input/soft-output decoding algorithms and this may follow the work presented in [2, 3]. These algorithms aim at minimising the probability of symbol or bit error and play a crucial part in iterative decoding.

In this paper, we focus on soft-combining techniques, which preserve the information obtained with each decoding attempt and incorporate this with retransmitted copies of a codeword. We are specifically interested in the post-decoding bit error probability when symbol-by-symbol MAP decoding is applied to linear block codes. For that purpose, we suggest a trellis-based decoding approach to be fit into the automata theory setting presented in [4, 5] rather than using generating functions [6, 7].

Moreover, this paper investigates two simple stop criteria to terminate retransmissions. The first criterion uses the sequence of signs of the soft-outputs at the MAP decoder and performs essentially a mapping of soft decisions onto hard decisions. The second approach for terminating retransmission exploits the cross-entropy between two subsequent retransmissions of a codeword which gives an indirect measure of the performance improvement that may be gained from the information contained in the latest retransmission.

The paper is organised as follows. Section 2 describes fundamentals of MAP decoding and introduces the principal soft-combining algorithm. In Section 3, the simple stop criterion, which uses the sign of the soft-output at the MAP decoder is presented. The stop criterion based on the cross-entropy between subsequent retransmissions is introduced in Section 4. Numerical examples are presented in Section 5. Finally, conclusions of the paper are given in Section 6.

2. Soft-combining algorithm

Let C denote an (n, k) block code over the Galois field $F = GF(2)$. The code C defines a one-to-one mapping of the k-dimensional information space F^k onto the n-dimensional vector space F^n. Let $\boldsymbol{u} = [u_1, u_2, \ldots, u_n] \in C$ be a codeword in the linear block code C and let $\boldsymbol{v} = [v_1, v_2, \ldots, v_n]$ denote a noisy observation at the output of a demodulator. In addition, we assume statistically independent source bits, systematic codes, and a memoryless channel. Eventually, the systematic linear block code C shall be defined by a parity check matrix $\mathbf{H}$. Then, the soft-output of a symbol-by-symbol MAP decoder for the

estimate $\hat{u}_i$ of the ith transmitted bit u_i is given by the log-likelihood ratio [7]:

$$L(\hat{u}_i) \triangleq L(u_i \mid \mathbf{v}) = \log \frac{P(u_i = 0 \mid \mathbf{v})}{P(u_i = 1 \mid \mathbf{v})} = \\ = L(u_i, v_i) + L_e(\hat{u}_i) \quad (1)$$

with joint log-likelihood ratio:

$$L(u_i, v_i) = \begin{cases} L(u_i) + L(v_i \mid u_i) & \text{for} \quad 1 \le i \le k, \\ L(v_i \mid u_i) & \text{for} \quad k < i \le n, \end{cases} \quad (2)$$

where $L(u_i)$ denotes the *a priori* value of the transmitted information bit u_i and $L(v_i \mid u_i)$ represents the soft-output value of the channel. The so-called extrinsic log-likelihood value $L_e(\hat{u}_i)$ is based on the indirect information about u_i due to the particular code in use and is given by

$$L_e(\hat{u}_i) = \log \left\{ \frac{\sum_{s=o}^{2^{n-k}-1} \prod_{\substack{j=1 \\ j \neq i}}^{n} \left(\frac{1-e^{-L(u_j \cdot v_j)}}{1+e^{-L(u_j \cdot v_j)}} \right)^{u^{\perp}_{s,j}}}{\sum_{s=0}^{2^{n-k}-1} (-1)^{u^{\perp}_{s,i}} \prod_{\substack{j=1 \\ j \neq i}}^{n} \left(\frac{1-e^{-L(u_j \cdot v_j)}}{1+e^{-L(u_j \cdot v_j)}} \right)^{u^{\perp}_{s,j}}} \right\} \quad (3)$$

with $u^{\perp}_{s,j}$ being the jth bit of the sth codeword $\mathbf{u}^{\perp}_s = \mathbf{s} \cdot \mathbf{H}$ in the dual code $C^{\perp}$ and $\mathbf{s} = bin(s)$ denotes the binary representation of the decimal number s. Note that the sign of the log-likelihood ratio $L(\hat{u}_i)$ provides a hard decision, i.e. $sign\{L(\hat{u}_i)\} = +1 \rightarrow \hat{u}_i = 0$ and $sign\{L(\hat{u}_i)\} = -1 \rightarrow \hat{u}_i = 1$ whereas the magnitude $abs\{L(\hat{u}_i)\}$ represents the reliability of the decision.

Since a MAP decoder processes soft-inputs and releases soft-outputs, the decoding outcome may be used along with a suitable ARQ scheme as the input for subsequent decoding attempts. To be more specific, a soft-combining algorithm can incorporate the information of each retransmission into the decoding procedure and this will be done on a symbol-by-symbol basis. The principal soft-combining procedure can be scheduled according to the following steps.

First transmission

1. Initialise $L^{(0)}(u_i)$ with *a priori* value of u_i.
2. Process soft-output value $L^{(1)}(v_i \mid u_i)$ of channel.
3. Compute extrinsic value $L_e^{(1)}(\hat{u}_i)$.
4. Compute log-likelihood ratio $L^{(1)}(\hat{u}_i)$.
5. Release decoded word if stop criterion is satisfied, otherwise continue with step 6.

Subsequent retransmissions

6. Request retransmission of codeword.
7. Use log-likelihood ratio $L^{(m-1)}(\hat{u}_i)$ obtained from $(m-1)$th retransmission as new *a priori* value of mth retransmission.
8. Process mth soft-output value $L^{(m)}(v_i \mid u_i)$ of channel.
9. Compute extrinsic value $L_e^{(m)}(\hat{u}_i)$ of mth retransmission.
10. Compute log-likelihood ratio $L^{(m)}(\hat{u}_i)$.
11. Release decoded word if stop criterion is satisfied, otherwise continue with step 6.

All the transmitted blocks may not need the same number of retransmissions to achieve reliable decoding. Some blocks may be detected with satisfactory reliability after the first decoding whereas some may need a few retransmissions. Also, some blocks would not even show any improvement in the post-decoding bit error rate (BER) performance with an increase in the number of retransmissions. Therefore, it is needed to observe whether improvement in the BER is possible to achieve, before requesting for another retransmission. If further performance improvement is not feasible, the retransmission of the same block can be stopped. This can be attained by specifying a criterion to terminate the retransmissions.

In previous publications, e.g. [8], availability of a genius observer has been assumed and was used to identify the undetectable errors. In this case, the genius observer requests a retransmission until a block is decoded error free or a specified maximum number of retransmissions is reached. Albeit this approach is rather idealistic, the performance characteristics obtained by employing a genius observer may serve as a benchmark for more realistic stop criteria.

3. Simple stop criterion based on hard decisions

A simple criterion for terminating retransmissions may be obtained by mapping the sequence of soft-outputs of the MAP decoder onto a sequence of hard decisions. A stop criterion can then be based on the hard decision for a complete word and may be defined as follows:

1. The hard decision for the complete word of the current transmission may be compared with that of any previous transmission and retransmissions will continue until a predefined number of those hard decisions, say three or four, are the same. This option requires sufficiently large storage space to buffer retransmissions until the examined word can finally be released.

2. Another option is to compare hard decisions of the current transmission only with that of the most recent retransmissions and terminate the soft-combining algorithm once they are the same. Since the reliability of a potential decision may be expected to increase with each retransmission, it is likely that subsequent decoding attempts result in the same hard decision.

The computer simulations have shown that the performance improvement obtained using the stop criterion based on the hard decisions of the current with any previous transmission is negligible compared to the results obtained based on

only the most recent transmission. Therefore, considering the complexity and storage requirements needed to compare all the previous hard decisions, only the most recent transmissions were considered for further study. This concept can also be extended to compare the current transmission with more than one previous retransmissions. However, with this criterion, at least two transmissions are required before stopping the retransmissions.

4. Stop criterion based on the cross-entropy

Moher has shown that cross-entropy provides a useful theoretical framework for iterative decoding [9]. Those ideas have been extended in [10] to show that cross-entropy can also be useful as a stop criterion for iterative decoding. Since the simple stop criterion introduced in Section 3 always needs a particular number of initial retransmissions, we follow the ideas presented in [9, 10] and exploit cross-entropy here between subsequent retransmissions as a more advanced stop criterion.

The cross-entropy between two subsequent retransmissions, which gives a measure of the closeness of two suitable distributions, is considered as a stop criterion in the decoding process. With this approach, the cross-entropy between the distributions corresponding to consecutive transmissions of the soft-combining process is estimated and compared to a threshold. Whenever the cross-entropy drops below the specified threshold, indicating a small change in the distribution from one retransmission to the next, the soft-combining algorithm is terminated.

The cross-entropy of two distributions $P(\hat{\mathbf{u}})$ and $Q(\hat{\mathbf{u}})$ over an alphabet F^n is defined as [9]

$$\boldsymbol{E}_P\left\{\log\frac{P(\hat{\mathbf{u}})}{Q(\hat{\mathbf{u}})}\right\}=\sum_{\hat{\mathbf{u}}\in F^n}P(\hat{\mathbf{u}})\log\frac{P(\hat{\mathbf{u}})}{Q(\hat{\mathbf{u}})}, \tag{4}$$

where $\boldsymbol{E}_p$ denotes the expectation operation over the distribution $P(\hat{\mathbf{u}})$ and F is the Galois field $GF(2)$. Assuming statistical independence of the symbol probabilities, we obtain

$$\log\frac{P(\hat{\mathbf{u}})}{Q(\hat{\mathbf{u}})}=\sum_i\log\frac{P(\hat{u}_i)}{Q(\hat{u}_i)}. \tag{5}$$

Let us assume that $P(\hat{\mathbf{u}})$ and $Q(\hat{\mathbf{u}})$ represent the distributions corresponding to the log-likelihood values of the $(m-1)$th and the mth retransmission, respectively. Then, the cross-entropy between the distributions corresponding to the $(m-1)$th and the mth retransmissions of the decoding process can be expressed as

$$\boldsymbol{E}_P\left\{\log\frac{P(\hat{\mathbf{u}})}{Q(\hat{\mathbf{u}})}\right\}=\sum_i\left[P(\hat{u}_i=0)\log\frac{P(\hat{u}_i=0)}{Q(\hat{u}_i=0)}+ +P(\hat{u}_i=1)\log\frac{P(\hat{u}_i=1)}{Q(\hat{u}_i=1)}\right]. \tag{6}$$

In Eq. (6), $P(\hat{u}_i)$ and $Q(\hat{u}_i)$ are the probabilities corresponding to the $(m-1)$th and the mth retransmission, respectively.

Using the log-likelihood value corresponding to the estimate $\hat{u}_i$ of the ith bit u_i in codeword $\mathbf{u}$, we can obtain the probabilities related to the $(m-1)$th retransmission as

$$P(\hat{u}_i=0)=\frac{e^{L^{(m-1)}(\hat{u}_i)}}{1+e^{L^{(m-1)}(\hat{u}_i)}}=\frac{1}{1+e^{-L^{(m-1)}(\hat{u}_i)}}, \tag{7}$$

$$P(\hat{u}_i=1)=\frac{1}{1+e^{L^{(m-1)}(\hat{u}_i)}}=\frac{e^{-L^{(m-1)}(\hat{u}_i)}}{1+e^{-L^{(m-1)}(\hat{u}_i)}}. \tag{8}$$

Here, $L^{(m-1)}(\hat{u}_i)$ represents the log-likelihood value for the estimate $\hat{u}_i$ which has been obtained after processing the information from the $(m-1)$th retransmission. Similarly, the probabilities related to the mth retransmission can be expressed as

$$Q(\hat{u}_i=0)=\frac{e^{L^{(m)}(\hat{u}_i)}}{1+e^{L^{(m)}(\hat{u}_i)}}=\frac{1}{1+e^{-L^{(m)}(\hat{u}_i)}}, \tag{9}$$

$$Q(\hat{u}_i=1)=\frac{1}{1+e^{L^{(m)}(\hat{u}_i)}}=\frac{e^{-L^{(m)}(\hat{u}_i)}}{1+e^{-L^{(m)}(\hat{u}_i)}}. \tag{10}$$

Using Eqs. (7) to (10), we obtain

$$P(\hat{u}_i{=}0)\log\frac{P(\hat{u}_i{=}0)}{Q(\hat{u}_i{=}0)}=\frac{1}{1+e^{-L^{(m-1)}(\hat{u}_i)}}\log\frac{1+e^{-L^{(m)}(\hat{u}_i)}}{1+e^{-L^{(m-1)}(\hat{u}_i)}}, \tag{11}$$

$$P(\hat{u}_i{=}1)\log\frac{P(\hat{u}_i{=}1)}{Q(\hat{u}_i{=}1)}=\frac{e^{-L^{(m-1)}(\hat{u}_i)}}{1+e^{-L^{(m-1)}(\hat{u}_i)}}\log\frac{1+e^{L^{(m)}(\hat{u}_i)}}{1+e^{L^{(m-1)}(\hat{u}_i)}}. \tag{12}$$

Let the difference between the two log-likelihood values corresponding to the two retransmissions be

$$\Delta L^{(m)}(\hat{u}_i)=L^{(m)}(\hat{u}_i)-L^{(m-1)}(\hat{u}_i). \tag{13}$$

Combining Eqs. (11), (12) and (13), we can express Eq. (6) as

$$\boldsymbol{E}_P\left\{\log\frac{P(\hat{\mathbf{u}})}{Q(\hat{\mathbf{u}})}\right\}=\sum_i\left[\log\frac{1+e^{-L^{(m)}(\hat{u}_i)}}{1+e^{-L^{(m-1)}(\hat{u}_i)}}+ +\frac{1}{1+e^{L^{(m-1)}(\hat{u}_i)}}\Delta L^{(m)}(\hat{u}_i)\right]. \tag{14}$$

With the proposed stop criterion, the request for a retransmission is terminated once the cross-entropy value as given in Eq. (14) falls below a specified threshold. In the present paper, however, the threshold has been set to a value of 10^{-3}.

5. Numerical example

In order to show the potential of performance improvement that can be gained from the soft-combining techniques, the computer simulations have been carried out considering a (7,4) Hamming code. Although the considered code is very simple, it can be used in constructing more powerful codes such as product codes. The examined (7,4) Hamming code can be defined by the parity check matrix:

$$\mathbf{H} = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}. \qquad (15)$$

Further, we assume that the source bits are equally likely to appear, i.e. $P(u_i = 0) = P(u_i = 1) = 0.5$ giving an *a priori* value of $L(u_i) = 0$. The Hamming code is used with binary phase shift keying (BPSK) over additive white Gaussian noise (AWGN) and Rayleigh fading channels. Soft-output of the channel $L(v_i \mid u_i)$ is specified by [11]:

$$L(v_i \mid u_i) = 4\frac{E_s}{N_0} a_i v_i . \qquad (16)$$

Here, a_i and v_i are the Rayleigh distributed fading amplitude and the ith output of a matched filter, respectively, and E_s/N_0 represents the signal-to-noise ratio (SNR). The Rayleigh fading channel is assumed being perfectly interleaved to ensure uncorrelated fading amplitudes a_i. Moreover, for an AWGN channel, amplitude $a_i = 1$.

The simulation results obtained for the (7,4) Hamming code using simple hard or soft decision decoding are compared with soft-combining on AWGN and perfectly interleaved Rayleigh fading channels as shown in Figs. 1 and 2, respectively. Since MAP decoding aims at minimising the symbol or bit error probability [2], performance has been evaluated in terms of post-decoding bit error rate. The obtained BER curves of this code using the proposed stop criteria are also included in Figs. 1 and 2. In the case

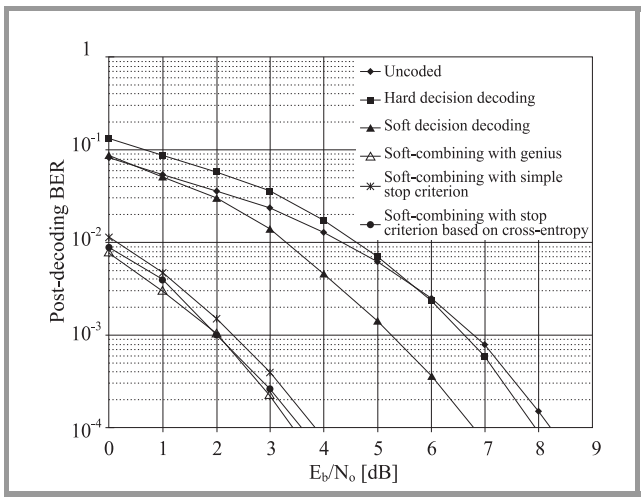


Fig. 1. BER performance of (7,4) Hamming code with soft-combining over AWGN channel.

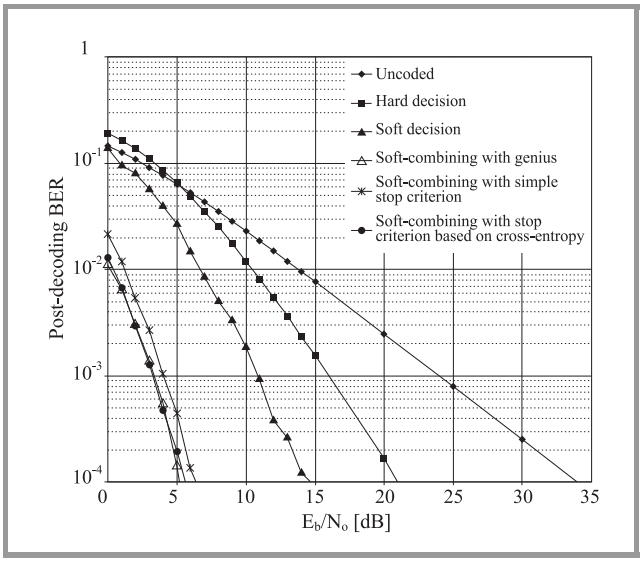


Fig. 2. BER performance of (7,4) Hamming code with soft-combining over Rayleigh channel.

of simple stop criterion, the hard decision of the current transmission is compared with that of the two most recent previous transmissions. Moreover, in these simulations, the maximum number of retransmissions has been set to 10.

It can be observed from the simulation results that error control coding using the (7,4) Hamming code provides a large performance improvement compared to an uncoded transmission, especially over fading channels. For example, the soft decision decoding of the (7,4) Hamming code gives coding gains of approximately 1.5 dB and 18 dB at a BER of 10^{-4} over AWGN and Rayleigh fading channels, respectively. Moreover, the soft-combining approach with a maximum of 10 retransmissions using the Hamming code over AWGN and Rayleigh fading channels provides additional gains of around 3 dB and 7 dB at the BER of 10^{-4}, respectively. It can also be observed that the proposed stop criteria perform very close to the one assuming a genius observer.

6. Conclusions

This paper presented soft-combining techniques for linear block codes using realistic stop criteria for retransmission termination. Two suitable stop criteria have been considered, firstly a very simple criterion based on hard decisions and secondly a more advanced criterion based on cross-entropy. It has been shown using an example that a significant performance improvement may be achieved with soft-combining techniques compared to simple soft decision decoding. The performance of the proposed stop criteria based on either the hard decisions of the current transmission with the two most recent retransmissions or the cross-entropy between the successive retransmissions was very close to the one assuming a genius observer over AWGN and Rayleigh fading channels. In addition, the stop

criterion based on cross-entropy has the advantage of requiring only two subsequent transmissions before terminating the retransmissions compared to at least three required for the one with hard decisions.

Acknowledgment

The authors wish to thank Christian Benien visiting from Aachen University of Technology, Germany, for optimising the MATLAB routines.

References

[1] D. Chase, "Code combining – a maximum-likelihood decoding approach for combining an arbitrary number of noisy packets", *IEEE Trans. Commun.*, vol. 33, no. 5, pp. 385–393, 1985.

[2] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimising symbol error rate", *IEEE Trans. Inform. Theory*, vol. 20, no. 3, pp. 284–287, 1974.

[3] C. R. P. Hartmann and L. D. Rudolph, "An optimum symbol-by-symbol decoding rule for linear codes", *IEEE Trans. Inform. Theory*, vol. 22, no. 6, pp. 514–517, 1976.

[4] H.-J. Zepernick and B. Rohani, "MAP decoding of LUEP codes based on a spectral domain analysis", *IEE Electron. Lett.*, vol. 36, no. 15, pp. 1292–1293, 2000.

[5] H.-J. Zepernick and B. Rohani, "On symbol-by-symbol MAP decoding of linear UEP codes", in *Proc. IEEE Globecom*, San Francisco, USA, Nov. 2000, pp. 1621–1626.

[6] G. Battail, M. C. Decouvelaere, and P. Godlewski, "Replication decoding", *IEEE Trans. Inform. Theory*, vol. 25, no. 3, pp. 332–345, 1979.

[7] E. Offer, *Soft-in/Soft-out Decoders for Linear Block Codes*. ITG-Fachbericht 130. Berlin: VDE-Verlag GmbH, 1994, pp. 31–40.

[8] H.-J. Zepernick and B. Rohani, "Soft-combining for linear unequal error protection codes", in *Proc. IEEE Int. Conf. Commun.*, Helsinki, Finland, June 2001, pp. 2970–2974.

[9] M. Moher, "Decoding via cross-entropy minimisation", in *Proc. IEEE Globecom*, Houston, USA, Dec. 1993, pp. 809–813.

[10] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes", *IEEE Trans. Inform. Theory*, vol. 42, no. 2, pp. 429–445, 1996.

[11] S. Riedel, "MAP decoding of convolutional codes using reciprocal dual codes", *IEEE Trans. Inform. Theory*, vol. 144, no. 3, pp. 1176–1187, 1998.

Manora Caldera received the B.Sc. (Eng.) degree in electronics and telecommunications with First Class Honours from the University of Moratuwa, Sri Lanka, in 1982. She then joined the Department of Electronic and Telecommunication Engineering at University of Moratuwa, Sri Lanka, as a lecturer. From 1984 to 1988, she carried out research at the Memorial University of Newfoundland, Canada, in the area of underwater communications and received her M.Eng. degree in 1986. She was with the School of Electrical and Computer Engineering at the Curtin University of Technology, Perth, Australia from 1992 to 2000 conducting research in the areas of mobile communications and error control coding. She received the Ph.D. degree from Curtin University of Technology in 2000. Currently, she is working as a Research Fellow at the Australian Telecommunications Cooperative Research Centre in the Wireless Program on coding and modulation topics. Her research interests include digital modulation, error control coding schemes, and mobile communication systems.
Australian Telecommunications Cooperative
Research Centre
GPO Box U1987, Perth WA 6845, Australia

Hans-Jürgen Zepernick – for biography, see this issue, p. 27.

Paper

Precise measurement of complex permittivity of materials for telecommunications devices

Takayuki Nakamura and Yoshio Nikawa

Abstract — In order to obtain precise complex permittivity of the dielectric materials obtained from the perturbation method a correction curve is made using the electromagnetic field simulator which applies transmission line modeling (TLM) method. In this experiment, generated microwave power with the frequency of 2.45 GHz is applied to heat dielectric material while measuring temperature dependence of complex permittivity of dielectric material. To obtain these objectives cavity resonator with cooling system is designed. It is found from the result that the accurate temperature dependence of complex permittivity of the materials can be obtained by the method presented here.

Keywords — perturbation method, TLM method, cavity resonator, simulation model, temperature dependence of complex permittivity, microwave measurement.

1. Introduction

It is known that the complex permittivity of materials usually changes depending on the frequency, on the temperature and on the compositions. Therefore, in designing or developing microwave devices, it is very important to study the temperature dependence of complex permittivity of materials over the wide temperature range. One of the usual techniques to obtain complex permittivity of materials is the perturbation method using the cavity resonator [1]. When the perturbation method is applied, it is necessary to satisfy two conditions. One of them is to use a small dielectric material compared with the volume of the cavity resonator, and the other one is that EM field distribution is not changed after inserting the dielectric materials into the cavity resonator. If dielectric material is fragile, it is very difficult to prepare a thin sample. Thus, EM field distribution is usually disturbed by the insertion of dielectric materials into the cavity. Therefore, calculation error by the perturbation method is accured according to the change of the EM field distribution changes, because the conditions of the perturbation method are not satisfied.

In this paper, correction curve is made by means of the EM field simulator, which applies the transmission line modeling method to reduce the error of complex permittivity of dielectric material obtained from the perturbation method. To measure the temperature dependence of complex permittivity of dielectric material and to heat the dielectric material simultaneously, microwave power with the frequency of 2.45 GHz is applied in this experiment, with use of a network analyzer and amplifier. To improve accuracy of measurements, only the dielectric materials have to be heated. Therefore, a rectangular cavity resonator with cavity cooling system was designed.

The theory and applications of TLM for EM field simulation are reviewed by Johns and Hoefer [3, 4]. The main advantage of the TLM method is to eliminate solving simultaneous numerous equations all over the structure. Therefore, the method consumes less computer memory and requires lower simulation time compared with the other simulators.

2. Theory

2.1. Perturbation theory

An example of cavity resonator is shown in Fig. 1. The dimensions of the cavity resonator in x, y and z directions are defined as a, b, and L, respectively.

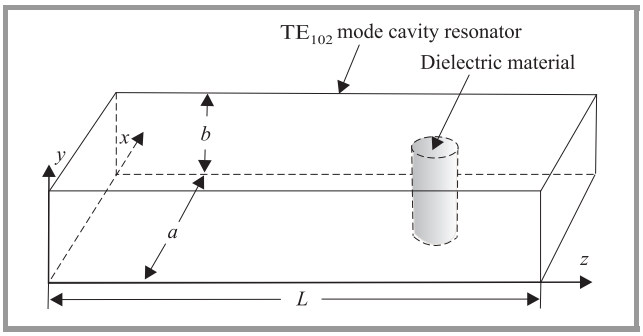


Fig. 1. Coordinates of cavity resonator.

If there is an electrical source with $J_e = j\omega\varepsilon_0(\varepsilon_r' - 1)\cdot E$, associated with a dielectric material inside the cavity, Maxwell's equations can be written as follows:

$$\nabla \times H = j\,\omega\,\varepsilon_0 E + J_e\,, \tag{1}$$

$$\nabla \times E = j\,\omega\,\mu_0 H\,, \tag{2}$$

where ω is the angular frequency, ε_0 and μ_o are permittivity and permeability of the free space, respectively, while E and H are the electric and the magnetic field, respectively.

If there is no dielectric material inside the cavity resonator, J_e equals to zero. For such a case the resonant angular frequency is defined as ω_0.
From Eqs. (1) and (2), Eq. (3) can be obtained as,

$$\int_V (E_0^* \cdot \nabla \times H + E \cdot \nabla \times H_0^* - H_0^* \cdot \nabla \times E + H \cdot \nabla \times E_0^*)\, dV = \\ = j(\omega - \omega_0) \int_V (\varepsilon_0 E_0^* \cdot E_0 + \mu_0 H_0^* \cdot H_0)\, dV + \int_{\Delta V} J_e E_0^*\, dV, \tag{3}$$

where V and ΔV are volume of the cavity and volume of the material. E_0^* and H_0^* are complex conjugates of electric and magnetic field intensities, ω and ω_0 are the angular frequencies before and after inserting dielectric materials into the cavity. From Gauss's theorem, we have $n \times E^* = n \times E = 0$. Hence Eq. (3) can be rewritten as follows:

$$\frac{\omega - \omega_0}{\omega_0} = j \frac{\int_{\Delta V} J_e E_0^*\, dV}{\omega_0 \int_V (\varepsilon_0 E_0^* \cdot E + \mu_0 H_0^* \cdot H)\, dV}. \tag{4}$$

The assumption to use the perturbation theory is that the EM field distribution will not be changed after inserting dielectric material into the cavity resonator. Under this assumption, E equals to E_0 and H equals to H_0. Thus Eq. (4) can be represented as follows:

$$\frac{\omega - \omega_0}{\omega_0} = -\frac{\int_{\Delta V} \{\varepsilon_0 (\varepsilon_r' - 1) |E_0|^2\}\, dV}{\int_V \{\varepsilon_0 |E_0|^2 + \mu_0 |H_0|^2\}\, dV}, \tag{5}$$

where ε_r' is the unknown relative complex permittivity to be obtained.
In this paper, we apply TE_{102} mode cavity. Therefore, only E_y, H_x and H_y components of the EM field can have non-zero values in the empty cavity. They are described as follows:

$$E_y = -C \frac{\omega \mu_0 k_x}{k_c^2} \sin(k_x x) \cos(k_y y) \sin(k_z z), \\ H_x = C \frac{j \beta_g k_x}{k_c^2} \sin(k_x x) \cos(k_y y) \cos(k_z z), \\ H_y = -jC \cos(k_x x) \cos(k_y y) \sin(k_z z), \tag{6}$$

where k_x, k_y and k_z, are $m\pi/a$, $n\pi/b$ and $p\pi/L$, respectively; β_g is $p\pi/L$, a, b and L are length of cavity resonator, and m, n and p are the mode indices in the cavity.
The complex resonant angular frequency of the cavity resonator ω is defined as follows:

$$\omega = \omega_r + j\omega_i, \\ \frac{\omega_i}{\omega_r} = \frac{1}{2 \cdot Q_L}, \tag{7}$$

where subscripts r and i stand for the real and the imaginary parts. Q_L stands for the loaded Q.

If complex resonant angular frequency is changed by insertion of the dielectric material into the cavity resonator, the left side part of Eq. (5) changes, resulting in:

$$\frac{\omega - \omega_0}{\omega_0} \approx \frac{\omega_r - \omega_{r0}}{\omega_{r0}} + \frac{j}{2} \left(\frac{1}{Q_L} - \frac{1}{Q_{L0}} \right). \tag{8}$$

From Eqs. (5) and (8), the complex permittivity can be obtained.

2.2. Simulation of resonant frequency and Q factor

With the perturbation method, it is assumed that the EM field distribution will not be changed after inserting the dielectric materials to the cavity. Therefore, the E_0 equals to E_0^* and the H_0 equals to H_0^* as adopted in Eq. (5). However, in the actual experiments, the EM field distribution is disturbed with insertion of the materials. As a result, the perturbation method always includes some error.
To simulate precise EM field and to obtain precise complex permittivity, a TLM simulator is applied to reduce measurement error as shown in Fig. 2.

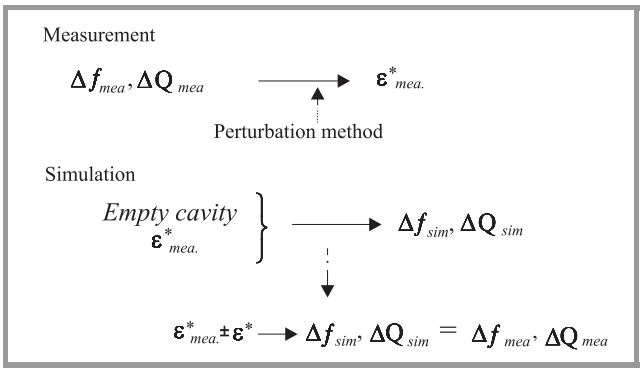


Fig. 2. Simulation method to reduce measurement error.

In Fig. 2, Δf_{mea}, ΔQ_{mea} and Δf_{sim}, ΔQ_{sim} represent the quantitative changes in resonant frequency and Q before and after inserting dielectric materials into the cavity both on the measurement and on the simulation, $\varepsilon^*_{mea.}$ and $\varepsilon^*_{sim.}$ are defined as complex permittivity obtained from the perturbation and simulation method, respectively.
At first, complex permittivity of the dielectric material is calculated by the perturbation method using measured resonant frequency and Q factor before and after inserting a dielectric material into the cavity resonator. Secondly, the empty rectangular cavity resonator was designed using TLM simulator (Micro-Strips, Fromerics). Then resonant frequency and Q are simulated using empty cavity resonator on the TLM method. Thirdly, the model of the rectangular cavity resonator with an inserted dielectric material is designed in the simulator. Then, resonant frequency and Q are

simulated by inputting characteristics of various complex permittivities of dielectric materials. After that, correction curve was made from calculated quantitative changes of resonant frequency and Q with and without placing the dielectric materials in the cavity resonator.

From these results, correction curves of the real and imaginary part of ε^* are obtained which are shown in Figs. 3 and 4, respectively.

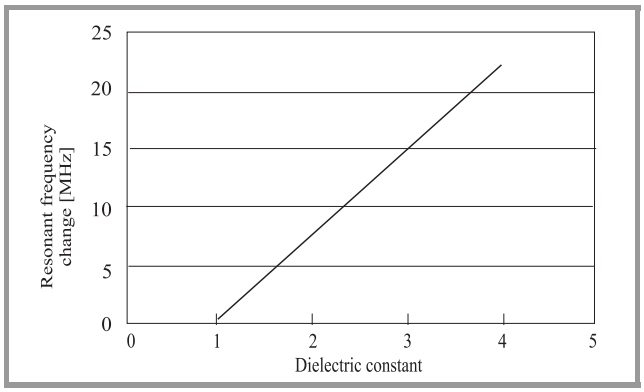


Fig. 3. Correction curve of real part.

Fourth, resonant frequency and Q are simulated with the adoption of the TLM simulator using the results of the complex permittivity on the perturbation method.

The calculated quantitative changes in resonant frequency and the Q for the TLM method are comparable with the result of quantitative changes in resonant frequency and Q for the perturbation method using correction curve. If these results are not fitting the correction curve, another value of complex permittivity will be input and the procedure will be repeated.

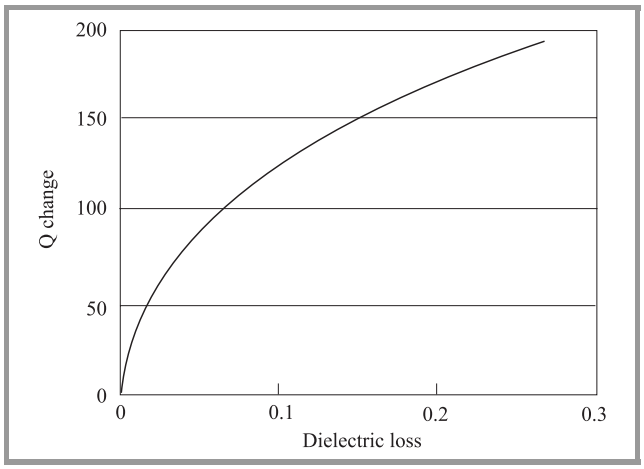


Fig. 4. Correction curve of imaginary part.

In the case, when the quantitative changes of resonant frequency and Q fits between the simulation and correction curves, the input parameter of complex permittivity is assumed to be a precise value.

3. Cavity resonator and measuring system

3.1. Experimental system

The TE_{102} mode rectangular cavity resonator is shown in Fig. 5. In this experiment, the length of the cavity is 147.2 mm to generate TE_{102} mode. The cross-sectional size is 110.0×27.0 mm^2. The coupling window and coupling disk are set at the optimum coupling position.

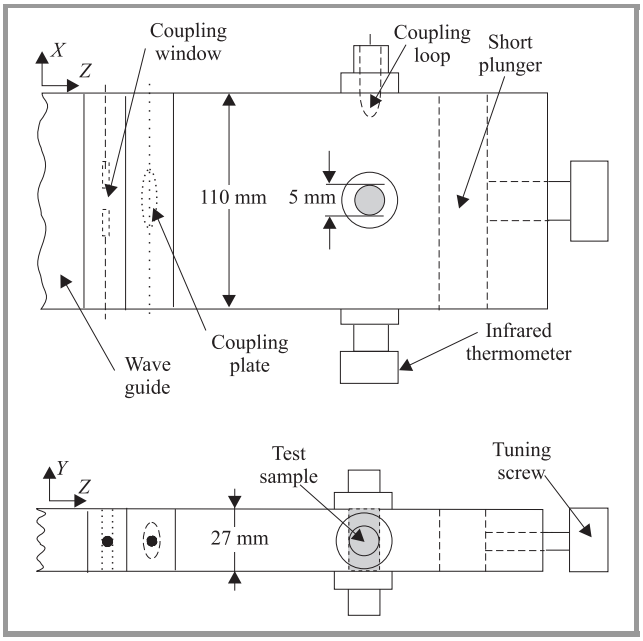


Fig. 5. TE_{102} mode rectangular cavity resonator.

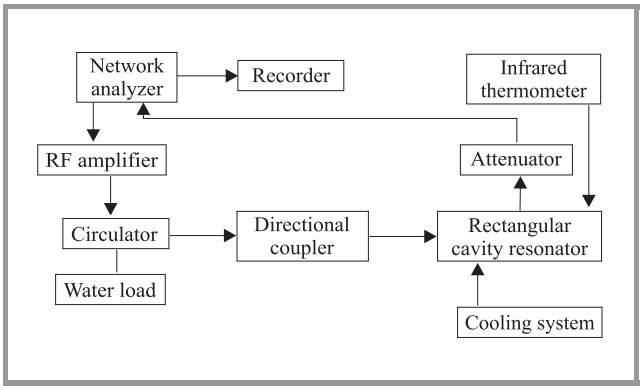


Fig. 6. Measuring system.

An infrared thermometer is used to measure surface temperature of the dielectric material from outside the cavity resonator through the small window on the cavity. The tuning screw can move the position of the short plunger to set initial resonant frequency [5, 6]. The dielectric material is inserted at the maximum EM field from the coupling window at 3/4 λ wavelength. Coupling loop to measure transmission power is inserted at the position of maximum magnetic field in the cavity.

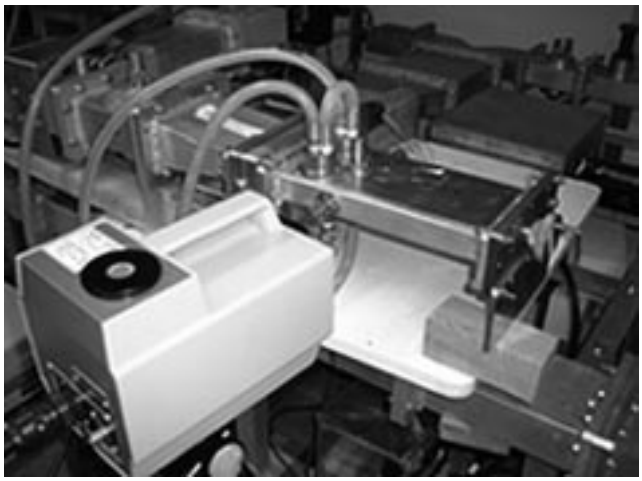


Fig. 7. Photograph of the measuring system.

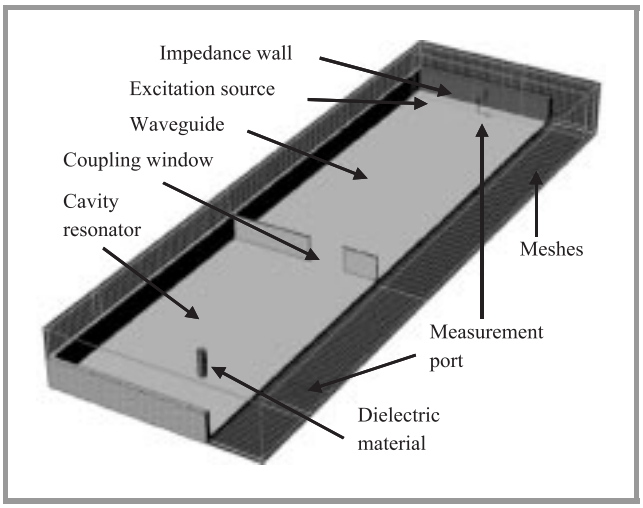


Fig. 8. Simulation model of cavity resonator.

The measuring system is shown in Figs. 6 and 7. Microwave power from the network analyzer (HP8753C) is amplified by means of a wideband high power amplifier (model A2325-5050-R R&K CO., LTD). An amplified microwave power is transmitted to the circulator, directional coupler and rectangular cavity resonator. The inserted coupling loop into the cavity resonator connects the network analyzer for measurement system transmitting power. A 30 dB attenuator, which can attenuate the equivalent power of amplifier gain, is set between rectangular cavity resonator and the network analyzer for protecting measuring circuit. The characteristics of resonant frequency and Q are measured by the network analyzer.

In the experiment, the center frequency of the signal fed to the cavity resonator was set at 2.45 GHz. Averaged duration time to tune of the resonant frequency was changed under the control of the sweeping span of the network analyzer. The temperature of dielectric material was measured by the infrared thermometer. In the experiment, little thick material was used intentionally for the evaluation of correction errors.

3.2. Simulation model of cavity resonator

A simulation model of cavity resonator is shown in Fig. 8. In the TLM method, the shape of cavity resonator is defined by metals, excitation source, the coupling loop, the dielectric material and the impedance wall. Excitation source and impedance wall are set at the waveguide side. Microwave power of TE_{10} mode is applies as excitation source. The impedance wall is used as the absorption wall and it has a defined optimum impedance. The waveguide and cavity resonator are separated by the coupling window. The dielectric material is set at maximum electric field at the 3/4 wavelengths from the coupling window. Measuring ports are set for calculation of the transmission property. Cavity was divided into an optimum number of cells. Especially, it was noted that around coupling window the dielectric materials and coupling loop are divided into smaller cells for increasing accuracy of the simulation.

4. Results

Real part of the complex permittivity versus temperature of polyethylene-terephthalate (PET) is shown in Fig. 9. It is clear that the results obtained using the TLM simulator are a little smaller than the reference measurement data. Nevertheless, it is apparent that the error is smaller than that using the perturbation method.

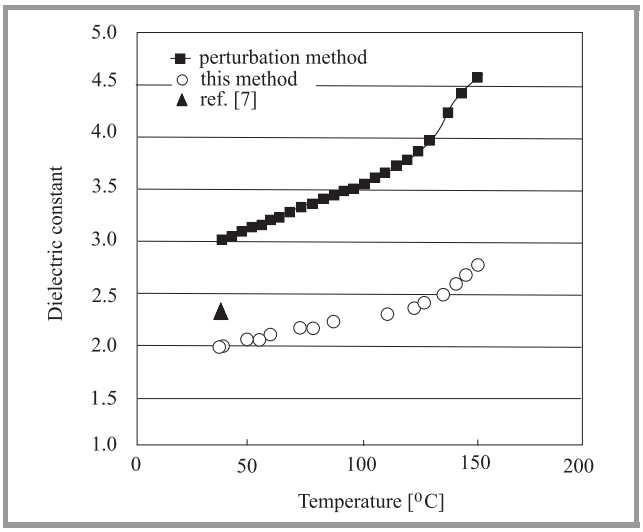


Fig. 9. Dielectric constant versus temperature of PET.

The results obtained from the perturbation method of dielectric loss versus temperature of PET are shown in Fig. 10 and the results obtained from TLM simulation and reference data are shown in Fig. 11. It is remarkable that the data obtained from TLM simulator almost fit to the reference data.

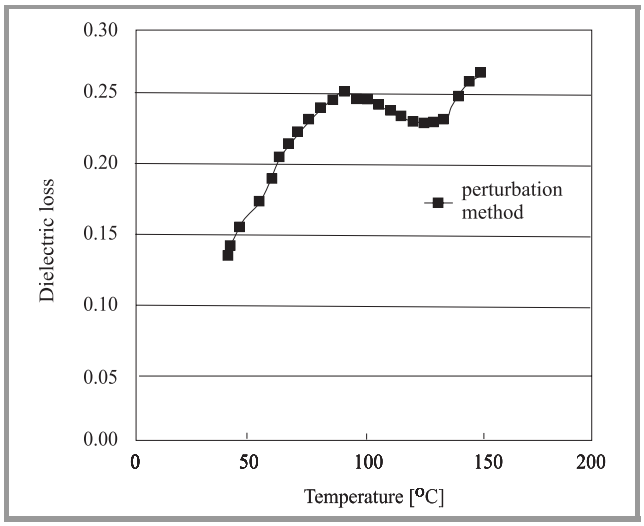


Fig. 10. Dielectric loss versus temperature of PET (perturbation method only).

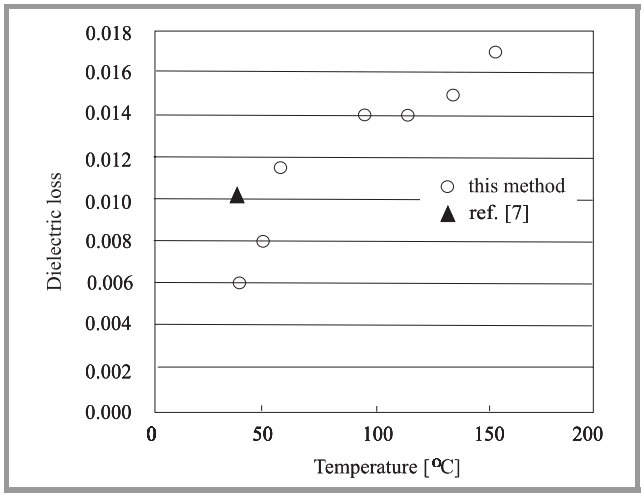


Fig. 11. Dielectric loss versus temperature of PET obtained by proposed method.

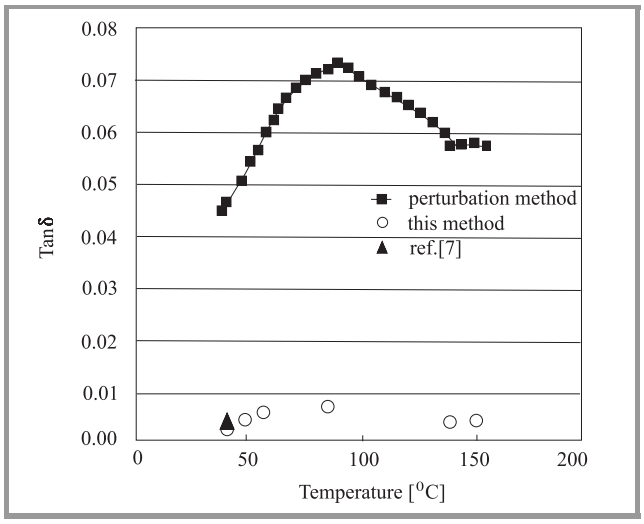


Fig. 12. Tanδ versus temperature of PET.

The loss tangent versus temperature of PET is shown in Fig. 12. In the figure, the simulation data is similar to the reference one obtained at the room temperature.
It is shown from these results that the TLM simulator can be used for the precise estimation of temperature dependence of complex permittivity.

5. Conclusions

The perturbation method is assumed to define that dielectric material is small enough to compare to the volume of the cavity resulting in EM field distribution not being changed after inserting the dielectric material into the cavity. But, for example, if we cannot use small dielectric material compared with the volume of the cavity because of the fragility of the material, the EM field distribution is changed by the insertion of dielectric material. Therefore, when perturbation method is applied, the correction error occurs because of the approximate calculation.
In this paper, the complex permittivity has been calculated using the TLM simulator. The designed correction curve obtained using the TLM simulator can reduce the calculation errors in complex permittivity of materials for the perturbation method. From these results, when TLM simulation and measurement with perturbation method are used together, the calculated data are expected to be similar to reference data.
In presented experiment, to measure temperature dependence of complex permittivity of the dielectric material, a microwave power with the frequency of 2.45 GHz was applied as measuring and heating power using a network analyzer and an amplifier. The averaged duration time to tune of the resonant frequency was changed by controlling of the sweeping span of the network analyzer. Thus stable temperature dependence of materials was measured. From these results, it is visible that the TLM method can be used to correct calculation error for the perturbation method.

Acknowledgment

The authors would like to thank to Prof. Fumiaki Okada, Department of Electrical Engineering of Kokushikan University for the usual advise about measurement method and to Mr. Hiroyuki Tanaka, Manufacturing Technology Company of SUMITOMO BAKELITE Group, for the useful suggestion to this project, and to Mr. Hiroshi Ando, Department of Mechanical Engineering of Kokushikan University, for preparing test samples.

References

[1] F. Okada, *Microwave Engineering – Principles & Applications*. Gakken-sha, 1993, pp. 369–370.

[2] P. B. Johns, "A symmetrical condensed node for the TLM method", *IEEE Trans. Microw. Theory Techn.*, vol. MTT-35, no. 4, pp. 370–377, 1987.

[3] P. B. Jhones and R. L. Beurle, "Numerical solution of 2-dimensional scattering problems using a transmission-line matrix", *Proc. Inst. Elec. Eng.*, vol. 118, no. 9, pp. 1203–1208, 1971.

[4] W. J. R. Hoefer, "The transmission-line method theory and applications", *IEEE Trans. Microw. Theory Techn.*, vol. MTT-33, no. 10, pp. 882–892, 1985.

[5] T. Nakamura, Y. Nikawa, and F. Okada, "Design of cavity resonator to measure temperature dependent complex permittivity of material", in *Proc. Korea-Japan Microw. Workshop* (KJMW), 2000, pp. 90–93.

[6] T. Nakamura, Y. Nikawa, and F. Okada, "Measurement of temperature depending microwave complex permittivity of resin", in *Proc. IEICE Electron. Soc.*, 2000, C-2-72, p. 99.

[7] T. Moreno, *Microwave Transmission Design Data*. Sperry Gyroscope Company, 1948, p. 206.

Takayuki Nakamura was born in Ibaraki, Japan, in August 2, 1972. He received the B.E. and M.E. degrees in electrical engineering from the Kokushikan University, Japan, in 1995 and 1997, respectively. He is currently studying for the Ph.D. degree at graduate school of Kokushikan University. His current research interests include measuring techniques in microwave field. Mr. Nakamura is a student member of the Institute of Electronics, Information and Communication Engineers (IEICE), Japan.
e-mail: g192102z@kokushikan.ac.jp
Graduate School of Engineering, Kokushikan University
4-28-1, Setagaya, Setagaya-ku, Tokyo, 154-8515 Japan

Yoshio Nikawa received the B.E., M.E. and Ph.D. degrees in electrical engineering form Keio University, Japan, in 1981, 1983, and 1986, respectively. In 1986, he was a Research Assistant at The National Defense Academy. From 1987 to 1988, he was the Invited Visiting Scholar at The University of Texas at Austin. He became an Associate Professor at The National Defense Academy in 1991. In April 1999, he joined Kokushikan University, Tokyo as a Professor in the Department of Electrical and Electronics Engineering. His research activities include microwave and millimeter-wave measurements and applications, electromagnetic scattering, propagation of electromagnetic waves, microwave and millimeter-wave heating and processing for medical and industrial applications. Dr. Nikawa is a member of the IEEE, Institute of Electronics Information, and Communication Engineers (IEICE), Japan, the Japan Society of Medical Electronics and Biological Engineering, and the Japanese Society of Hyperthermic Oncology.
e-mail: nikawa@kokushikan.ac.jp
Graduate School of Engineering, Kokushikan University
4-28-1, Setagaya, Setagaya-ku, Tokyo, 154-8515 Japan

Regular paper

Method of analytical regularization based on the static part inversion in wave scattering by imperfect thin screens

Alexander I. Nosich

Abstract — The paper is focussed on the development of the method of analytical regularization (MAR) in electromagnetic wave scattering and absorption by imperfect scatterers shaped as thin screens.

Keywords — wave scattering, imperfect boundary conditions, thin screens, regularization.

1. Imperfect boundary conditions

As known (see [1 – 6]), if the thickness of imperfect scatterer is small compared to the free-space wavelength, the wave scattering problem can be simplified to exclude the internal field from consideration. This is done by assuming the scatterer thickness to be zero but at the same time introducing specific boundary conditions modified with respect to the perfectly electric conducting (PEC) boundary conditions. In these conditions, certain effective parameters appear, accumulating the values of the thickness and material constants of the scatterer. These parameters couple together the limit values of the tangential field components $\vec{E}_T^{\pm}$ and $\vec{H}_T^{\pm}$ on the two sides of the scatterer, namely, their jumps and their mean values. In all, there are three different conditions of this type [4]:

1) resistive:

$$\frac{1}{2}[\vec{E}_T^+ + \vec{E}_T^-] = Z_0 R \vec{n} \times [\vec{H}_T^+ - \vec{H}_T^-], \qquad (1)$$

$$\vec{E}_T^+ = \vec{E}_T^-, \qquad (2)$$

2) material, or magneto-dielectric:

$$\frac{1}{2}[\vec{E}_T^+ + \vec{E}_T^-] = Z_0 R \vec{n} \times [\vec{H}_T^+ - \vec{H}_T^-], \qquad (3)$$

$$\frac{1}{2}[\vec{H}_T^+ + \vec{H}_T^-] = -\frac{S}{Z_0}\vec{n} \times [\vec{E}_T^+ - \vec{E}_T^-], \qquad (4)$$

and

3) impedance one:

$$\frac{1}{2}[\vec{E}_T^+ + \vec{E}_T^-] = Z_0 R \vec{n} \times [\vec{H}_T^+ - \vec{H}_T^-] + W[\vec{E}_T^+ - \vec{E}_T^-], \qquad (5)$$

$$\frac{1}{2}[\vec{H}_T^+ + \vec{H}_T^-] = -\frac{S}{Z_0}\vec{n} \times [\vec{E}_T^+ - \vec{E}_T^-] - W[\vec{H}_T^+ - \vec{H}_T^-]. \qquad (6)$$

The first condition appears in the case of a thinner-than-skindepth metal scatterer of finite conductivity; the second one is derived for a thin magneto-dielectric scatterer (material), and the third one for a PEC scatterer covered with a thin material layer. Thin scatterers are commonly called screens and can be flat or curved. Of the three mentioned types of screens, the first two (resistive and material) are partially transparent while the third one (impedance) is non-transparent. Note that conditions (5) and (6) are of the most general form, and the first two ones can be viewed as particular cases of that form.

So, any mentioned type of thin scatterers can be simulated by using at most three effective parameters. There exists an ambiguous terminology about these parameters. We shall use the one proposed in [4] and call them resistivities: electric R, magnetic S, and so-called cross-resistivity W. For example, in the case of a thinner-than-skindepth metal sheet and a thin low-contrast dielectric sheet having normalized material parameters ε_r and μ_r, respectively [2, 4, 5]:

$$R_m = (Z_0 b\sigma)^{-1}, \qquad R_d = i[k_0 b(\varepsilon_r - 1)]^{-1} \qquad (7)$$

while $S = i\infty$ and $W = 0$. Here b is the thickness, σ is the conductivity, Z_0 is the free-space impedance, k_0 is the free-space wavenumber, and it is assumed that $\mu_r = 1$, $|\varepsilon_r - 1| \ll 1$ and $k_0 b \ll 1$. In the case of a thin single-layer high-contrast material sheet, it can be shown (see [1, 3, 4]) that

$$R = Z^2 S = \frac{i}{2} Z \cot\left(\frac{1}{2}\varepsilon_r^{1/2}\mu_r^{1/2} k_0 b\right), \qquad W = 0, \qquad (8)$$

where $Z = (\mu_r/\varepsilon_r)^{1/2}$, and it is assumed that $|\varepsilon_r \mu_r| \gg 1$ and $k_0 b \ll 1$. Note that for a multi-layer sandwich-like material sheet, $W \neq 0$ [4].

It is interesting to recall that the conditions (1), (2) and (3), (4) were at first proposed empirically. They had been in extensive use for quite a long time (e.g., see [7 – 10]) before being fully grounded in [2] and [3], respectively. Mathematically rigorous derivation of the expressions (7) and (8) for the resistivities was also done in [2, 3] and confirmed more relaxed derivations of [1, 4 – 6].

For the sake of completeness, it should be mentioned that material screens have been also simulated by introduc-

ing the so-called "higher-order" imperfect boundary conditions [5]. The latter involve not only the limit values of tangential fields but also their normal derivatives. For example, condition (4) is modified to take the following "first-order" form (see [11]):

$$\frac{1}{2}[\vec{H}_T^+ + \vec{H}_T^-] = -\frac{1}{Z_0}\vec{n}\times\left(S - \frac{\partial}{k_0\partial n}\right)[\vec{E}_T^+ - \vec{E}_T^-]. \quad (9)$$

In this paper, we shall base our considerations on the "zeroth-order" boundary conditions $(1)-(6)$ although one can extend MAR to the conditions of [11]. Note that conditions of the order higher than 1 are ambiguous from the viewpoint of the solution uniqueness [12]. Further, although conditions $(1)-(6)$ are derived for infinite thin sheets, we shall apply them to the limited screens like strips and disks, which have sharp edges. We shall not discuss here the validity of the modified conditions near the edges, and only remark that, according to [2], this is less ambigous than using together the PEC conditions and the zero thickness of a screen with edges. Instead, we shall imply that each time some appropriate edge condition is included in the problem formulation, that guarantees the solution uniqueness.

Impenetrable thin imperfect screen is a special case when three resistivities are not independent. As shown in [4], they should satisfy the relationship

$$4(RS + W^2) = 1. \quad (10)$$

Therefore, in this case one of three resistivities can be eliminated, and the pair of conditions (5) and (6) can be equivalently formulated in terms of two other effective parameters: surface impedances Z^+ and Z^-, i.e., as the two-side Leontovich boundary conditions $[4-6]$:

$$\vec{E}_T^{\pm} \mp Z_0 Z^{\pm}\vec{n}\times\vec{H}_T^{\pm} = 0. \quad (11)$$

Equations (11) are equivalent to (5) and (6) provided that

$$R = \frac{Z^+Z^-}{Z^+ + Z^-},\ S = \frac{1}{Z^+ + Z^-},\ W = \frac{1}{2}\frac{Z^+ - Z^-}{Z^+ + Z^-},\ Z^+ + Z^- \neq 0. \quad (12)$$

For example, a thin PEC screen coated with different layers of magneto-dielectrics of parameters $b^{\pm}$, $\varepsilon_r^{\pm}$ and $\mu_r^{\pm}$, has the values of the surface impedances given by $[4-6]$

$$Z^{\pm} = -iZ\tan[(\varepsilon_r^{\pm}\mu_r^{\pm})^{1/2}k_0 b^{\pm}]. \quad (13)$$

It is important to note that if any material parameter of an imperfect screen is not real but complex-valued, then the resistivities obtain non-zero real parts, which are responsible for the dissipation losses. Therefore, by modifying the boundary conditions, one can study not only the wave scattering but also the wave absorption. Besides, if any of the quantities ε_r, μ_r or h varies along the screen surface, then the resistivities R, S, W are the functions of coordinates.

2. About the method of analytical regularization

Hence, there arises a challenge to extend or modify the previously existed analytical and numerical solutions of the PEC-screen wave scattering problems to the three mentioned types of imperfect thin screens. In computational electromagnetics, one of the most powerful and efficient approaches is based on the integral equations (IE) – see [13]. Here, the method of analytical regularization, i.e., a semi-inversion of the full-wave singular IE [14], is the one that guarantees numerical convergence. General scheme of MAR works as follows. Commonly, the boundary PEC conditions generate a singular IE of the first kind: $\hat{G}X = F$. Split the operator $\hat{G}$ into two parts, $\hat{G}_1$ and $\hat{G}_2$. Provided that the former has a known inverse $\hat{G}_1^{-1}$, the original equation can be converted to the second-kind one: $X + \hat{A}X = B$, where $\hat{A} = \hat{G}_1^{-1}\hat{G}_2$ and $B = \hat{G}_1^{-1}F$. However, this scheme is mathematically justified only if the operator $\hat{A}$ is compact, i.e., its norm $||\hat{A}|| < \infty$ in the functional space L_2. This implies inherently that the inverted operator $\hat{G}_1$ must be a singular one while $\hat{G}_2$ is regular. It is possible to point out several different ways of extracting out an invertible singular part of original equation. It corresponds to either canonical-shape or to the high-frequency or to the static part of the full-wave IE operator [14]. Once this has been done, it is guaranteed, thanks to the Fredholm theorems, that the usual discretization schemes converge to the exact solution in the point-wise sense. Here, the convergence is understood as a possibility to minimize the computation error to machine precision by solving progressively larger matrices.

The variety of problems solved by MAR with the static part inversion covers a wide class of PEC zero-thickness screens [14]. Among them there are single strips and strip gratings, strip irises in a waveguide, periodic circular waveguides, open circular cylindrical screens and collections of them, longitudinally slitted infinite cone, circular disk, spherical cap, finite circular hollow pipe, etc. Any of the listed problems is reduced first to a single singular IE or a coupled pair of IEs of the first kind. A limit form of IE, corresponding to the static problem, can be inverted analytically based on the theory of the Cauchy integrals. Application of this result to the full-wave IE leads to an IE of the Fredholm 2nd kind with a smooth kernel. Hence, the existence of the unique solution is guaranteed. Numerical solution is then easy to obtain by using any reasonable discretization scheme, and the validity of the matrix truncation is justified.

Discretization and partial inversion can be joined together in a single procedure, if one uses the set of orthogonal eigenfunctions of the static part of the full-wave operator as a projection basis in a Galerkin scheme [14]. As the accuracy of computations is improved by increasing the number of equations and is limited only by the digital precision of computer used, MAR may be called a "numerically exact" approach. The number of equations needed for a practi-

cal accuracy of 3–4 digits is normally slightly greater than electrical size of a PEC scatterer. As we shall see, in the 2D case of the H-polarization, these solutions can be directly extended to resistive scatterers as well, because non-zero R does not change the static limit of IE. In the E-polarization case, the situation is different. Here, non-zero R changes the static behavior of the scatterer. However, in the E-case, second-kind IE obtained from the imperfect boundary condition is already a Fredholm one, whose operator vanishes in the static limit. For the axisymmetric 3D screens, the same is valid with respect to IEs for two potential functions, which are frequently taken as E_ϕ and H_ϕ. In each case, the obtained second-kind equations correspond to the static-part inversion. It is also necessary to note that if there exists a one-to-one mapping $X = CZ$, $Z = C^{-1}X$, then one can build a MAR analysis on the discretization of the equivalent operator equation $Z + (C^{-1}AC)Z = C^{-1}B$, which is also a Fredholm second-kind one. For example, operators C and C^{-1} can be direct and inverse integral Fourier transforms (in the single strip scattering) or discrete Fourier transforms (in the strip grating scattering) or integral Hankel transforms (in the disk scattering).

3. Resistive strip scattering

Consider an example in 2D: the scattering of a given time-harmonic ($\sim e^{-i\omega t}$) electromagnetic field by a resistive strip [9, 15, 16], whose contour of cross-section is an open curve M in the plane (x,y). Here, two alternative polarization cases can be treated separately. Generally, the scattered field has to satisfy the Helmholtz equation off M, boundary conditions on M, edge condition near the sharp edges of the screen, and the radiation condition at infinity. In the case of the H-polarization, the role of potential function is played by the magnetic-field component parallel to the strip generatrix, $H = H_z$. Then, the boundary conditions (1), (2) take the form as

$$\frac{1}{2}\left(\frac{\partial H^+}{\partial n} + \frac{\partial H^-}{\partial n}\right) + ik_0 Z_0 R(H^+ - H^-) = 0, \qquad (14)$$

$$\partial H^+/\partial n - \partial H^-/\partial n = 0. \qquad (15)$$

After decomposing the total field into the sum of the incident H^{in} and scattered one H^{sc} and presenting the latter in the form of a double-layer potential, one obtains a hypersingular integral equation of the second kind:

$$ik_0 RX(\vec{r}) + \frac{\partial}{\partial n}\int_M X(\vec{r}\,')\frac{\partial}{\partial n'}G_0(\vec{r},\vec{r}\,')d\vec{r}\,' = -\frac{\partial H^{in}(\vec{r})}{\partial n}, \qquad (16)$$

where $G_0 = i/4H_0^{(1)}(k|\vec{r} - \vec{r}\,'|)$ is the 2D free-space Green's function ($H_0^{(1)}$ stands for the Hankel function), and $X = H^+ - H^-$ is the unknown surface-current density.

Note that in Eq. (14) and hence in IE (16), the term containing the product $k_0 R$ is a simple perturbation to the PEC boundary condition and the IE, respectively. That is why analytical regularization of IE (16) can be done in the same way as for a PEC screen, and a smooth passing to the limits $k_0 \to 0$ and $R \to 0$ is possible at every step of this procedure. The inversion of the static part of IE (16) is based on the diagonalization of the integral operator with a hyper-type singularity. This is due to the existence of a set of orthogonal eigenfunctions of the IE static limit: for example, if M is a straight interval, i.e., if the strip is flat, they are the weighted Chebyshev polynomials. Further details of this analysis can be found in [15, 16], for a circularly curved resistive strip and for a flat resistive strip, respectively.

In the alternative case of the E-polarization, the role of potential function is played by the electric-field component parallel to the strip generatrix, $E = E_z$. Then the resistive boundary conditions (1), (2) take the form as

$$\frac{1}{2}\left(\frac{\partial E^+}{\partial n} - \frac{\partial E^-}{\partial n}\right) + \frac{ik_0}{Z_0 R}(E^+ + E^-) = 0, \qquad (17)$$

$$E^+ - E^- = 0. \qquad (18)$$

These conditions, together with the single-layer representation of the scattered field E^{sc}, lead us to the following IE of the second kind:

$$Y(\vec{r}) + \frac{ik_0}{R}\int_M Y(\vec{r}\,')G_0(\vec{r},\vec{r}\,')d\vec{r}\,' = -\frac{ik_0}{R}E^{in}(\vec{r}), \qquad (19)$$

where $Y = \partial E^+/\partial n - \partial E^-/\partial n$ is the unknown surface-current density.

Integral equation (19) has a logarithmic-singular kernel G_0. Such a singularity is integrable, hence this IE is of the Fredholm second-kind provided that $R \neq 0$. That is why it can be discretized by using any usual discretization scheme with local or global-basis expansion functions. One can see that the norm of this IE operator is finite for any $R \neq 0$ and proportional to k_0. Therefore, it may be stated that IE (19) is based on the analytical inversion of the static limit of the full-wave scattering problem. However, unlike in the H-case, the limit forms of (17), (19) for $k_0 \to 0$ and $R \to 0$ are essentially different. For any $R \neq 0$ the static limit of (17) is not the PEC condition; besides, the static limit of the solution to IE (19) is identical zero. Still besides, one can see that if R is purely imaginary or purely real, the ratio $ik_0^{-1}R$ plays the role of a Lavrentyev or a Bakushinsky regularization parameter, respectively [17, 18].

In the case of a circularly curved open resistive strip of radius a (Fig. 1), IEs (16) and (19) can be transformed with discrete Fourier transform, and further static-part inversion can be done in the transform domain [15]. The radar cross-sections of such a strip illuminated by a plane wave are presented in Figs. 2 and 3. In the flat-strip case, IEs can be transformed into the integral Fourier transform domain. Static part inversion can be done in the latter

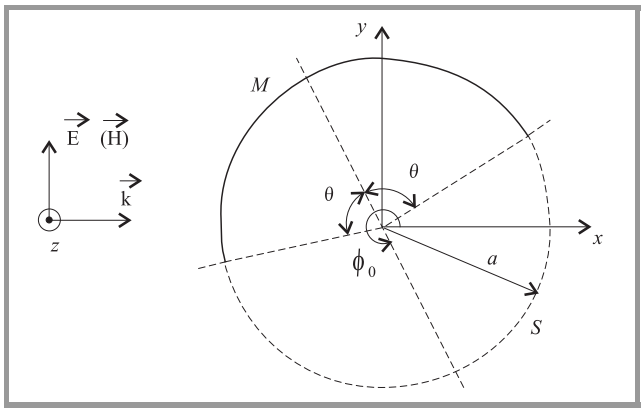


Fig. 1. Scattering geometry of the cross-section of a circularly curved resistive strip.

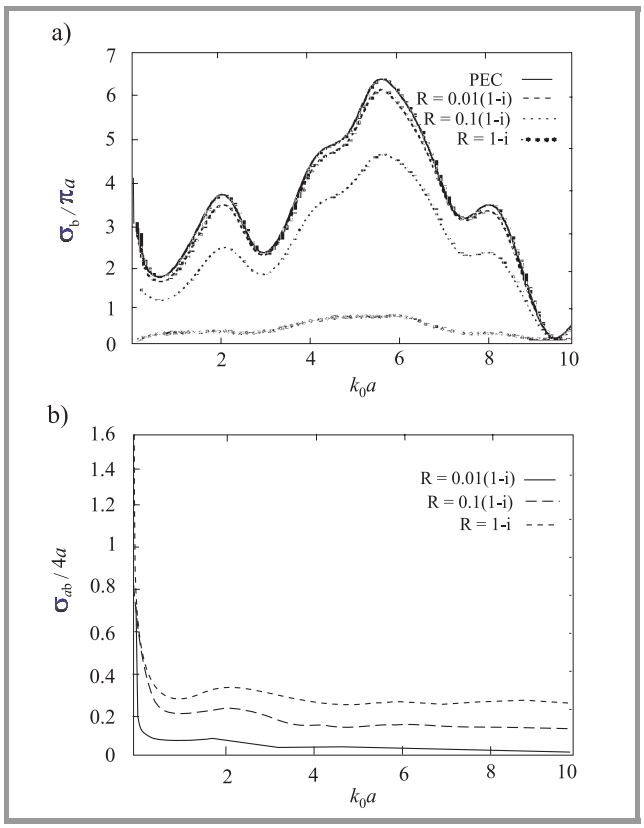


Fig. 2. Lossy resistive strip characteristics in the E-polarization: (a) normalized backward (radar) scattering cross-sections, and (b) absorption cross-sections as functions of the normalized frequency k_0a. $\phi_0 = 170^o$, $\theta = 90^o$.

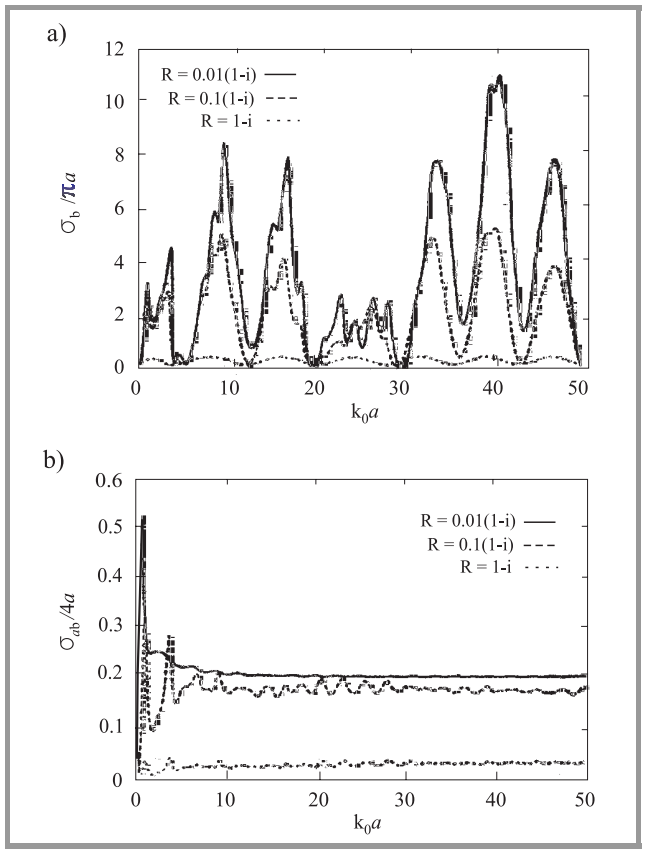


Fig. 3. Lossy resistive strip characteristics in the H-polarization: (a) normalized backward (radar) scattering cross-sections, and (b) absorption cross-sections as functions of the normalized frequency k_0a. $\phi_0 = 170^o$, $\theta = 90^o$.

domain as well, (see [16]), with the Bessel functions (transformed Chebyshev polynomials) as a basis. The algorithms based on the space-domain and transform-domain MAR are equally high-efficient. A final remark can be done about the extension of MAR solution to a strip with the resistivity varying along the contour M. In this case, R in IEs (16) and (19) must be viewed as a function of $\vec{r}$. This circumstance does not change the basic properties of IEs, and hence the same MAR schemes work out, although the rate of convergence gets worse. Based on such a modified algorithm, a cylindrical reflector antenna with a variable-resistivity edge loading was analyzed in [19].

4. Material and impedance strip scattering

In the homogeneous material-strip scattering, we start from the boundary conditions (3), (4). Together with the representation of the scattered field as a sum of a single- and a double-layer potentials, they yield now not one but two second-kind decoupled IEs. Each of the latter can be further treated in the same way as it was done previously for a resistive strip scattering with a H-wave and an E-wave incident, respectively. Paper [20] dealt with solving these equations after converting them to the integral Fourier-transform domain. In the E-polarization case, material thin-strip boundary conditions lead us to a similar pair of the second-kind decoupled IEs, with the parameters R and S interchanged. Hence, in the material-strip scattering, a difference in the electromagnetic behavior between the E-wave and H-wave cases vanishes. Note that, to build the scattered field, the contributions from the solutions of the both IEs

must be taken into account. In the scattering by a multilayer material strip or an impedance strip, the boundary conditions (5), (6) or, equivalently, Eqs. (11) should be used. They bring us to the pair of coupled IEs of the second kind, where cross-resistivity W plays the role of the coupling parameter. If the surface impedances are the same: $Z^+ = Z^-$, then $R = \frac{1}{2}Z^+$, $S = (2Z^+)^{-1}$, $W = 0$, and hence the IEs again decouple. Special case of $Z^+ = -Z^-$ can be considered as well [4].

5. Imperfect strip grating scattering

In the case of the scattering of plane waves by a flat grating made of identic periodically spaced resistive, material or impedance strips of period l (Fig. 4), the same MAR approach as for a single strip can be used. This is due to the fact that in the kernels of corresponding IEs, the quasi-periodic free-space Green's function, G_p, takes the place of G_0. As $G_p = G_0 + P$, where P is a regular function at $\vec{r} \to \vec{r}'$, the singularities are kept the same, and hence a similar static-part inversion results in the regularized infinite-matrix equations. A grating of flat resistive

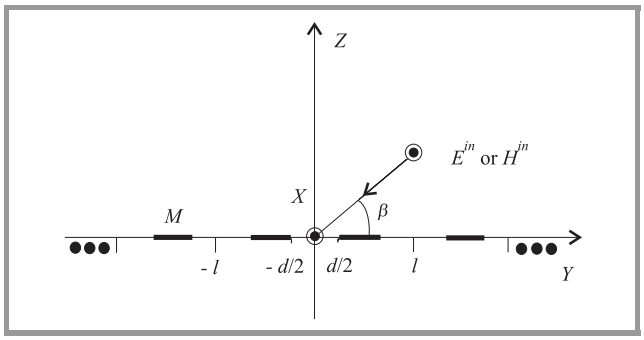


Fig. 4. Scattering geometry of the cross-section of a flat resistive-strip or dielectric-strip grating.

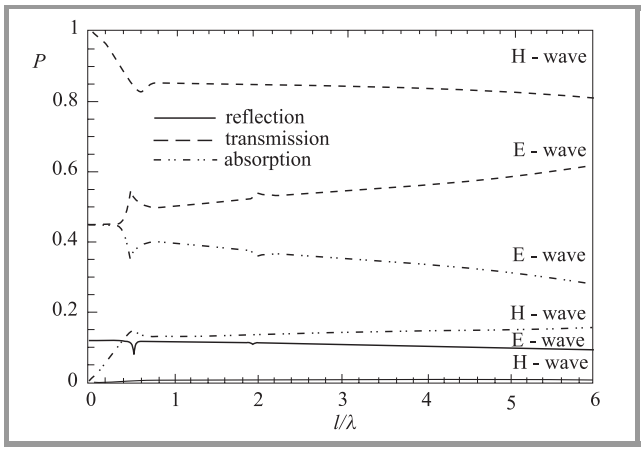


Fig. 5. Transmitted, reflected and absorbed power fractions for the scattering by a resistive-strip grating as functions of the normalized frequency l/λ ($k_0 = 2\pi/\lambda$). $\beta = 30°$, $d/l = 0.5$, $R = 1$.

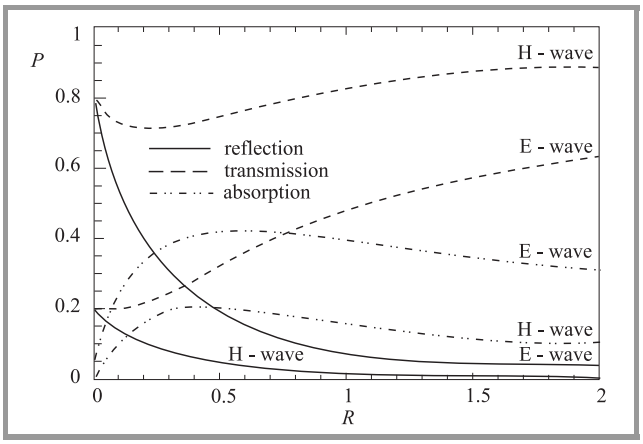


Fig. 6. Power fractions for the scattering by a resistive-strip grating as functions of the normalized resistivity. $\beta = 30°$, $l/\lambda = 0.5$, $d/l = 0.5$.

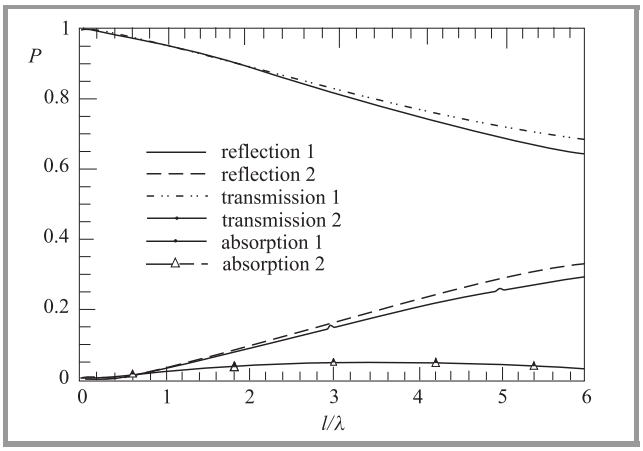


Fig. 7. Transmitted, reflected and absorbed power fractions for the H-wave scattering by a dielectric-strip grating as functions of the normalized frequency for $\beta = 90°$, $d/l = 0.5$, $b/l = 0.01$, $\varepsilon_r = 10 + i$, $\mu_r = 1$. Two models are compared: 1 – high-contrast, i.e. R and S based on Eq. (8), and 2 – low-contrast, i.e. R based on Eq. (7) and $S = i\infty$.

strips has been considered in [10, 21 – 24]. Note that in the H-wave case, the results published in [21] were erroneous as the matrix elements did not decrease with greater indices. This is because no regularization of a hyper-singular IE was performed. The latter was developed in [22, 23], where it was noted that the same mistake was characteristic for the other papers considering the scattering of H-waves by a resistive strip grating. In the E-wave case, the results published in [10, 21] are correct and agree with the data of [22 – 24] where different discretizations were used. In Figs. 5 and 6, the power fractions (normalized powers of transmission, reflection and absorption) in the plane wave scattering by a resistive strip grating are presented [22]. In Figs. 7 and 8, the same is presented for the scattering by a thin-strip dielectric grating, which was also considered in [22]. Note that the latter results were confronted with the exact solution of the Fredholm second-kind

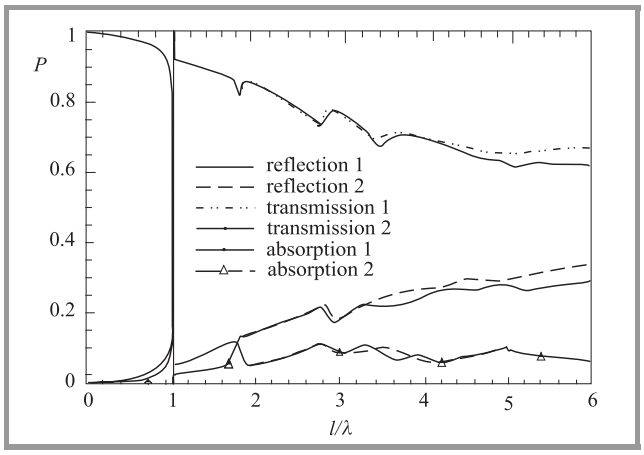


Fig. 8. Transmitted, reflected and absorbed power fractions for the E-wave scattering by a dielectric-strip grating as functions of the normalized frequency for $\beta = 90°$, $d/l = 0.5$, $b/l = 0.01$, $\varepsilon_r = 10 + i$, $\mu_r = 1$. Two models are compared: 1 – high-contrast, i.e. R and S based on Eq. (8), and 2 – low-contrast, i.e. R based on Eq. (7) and $S = i\infty$.

"domain" IE for a finite-thickness dielectric strip grating. Comparison showed a very good agreement for the strip thickness being 1/10 of the period. The scattering by an impedance strip grating has not been analyzed with MAR so far.

6. Imperfect disk scattering

Consider a curved rotationally symmetric disk supporting resistive boundary conditions (1), (2). Here, two simplest diffraction problems arise: excitation of the disk by either a coaxial vertical electric dipole (CVED) or a magnetic one (CVMD). Figure 9 demonstrates an example of such a scattering geometry associated with a flat disk of radius a placed on top of a dielectric substrate. The field in such a geometry is ϕ-independent and can be expressed via a single potential function: H_ϕ or E_ϕ, respectively [25]. In the case of electric (or magnetic) dipole, we arrive at the singular IE of the second kind similar to IE (16) (or (19)),

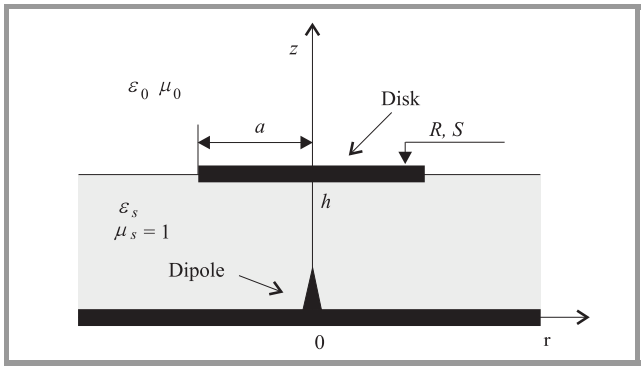


Fig. 9. Scattering geometry of the cross-section of a circular disk on a grounded dielectric substrate.

with the ϕ-independent scalar 3D Green's function K_0 taking place of G_0:

$$K_0(\vec{r},\vec{r}') = r' \int_0^{2\pi} \frac{e^{ik_0\rho}}{\rho} \cos\psi d\psi, \tag{20}$$

where

$$\psi = \phi - \phi', \quad \rho = [r^2 + r'^2 + 2rr'\cos\psi + (z-z')^2]^{1/2}. \tag{21}$$

The domain of integration in IE correspondingly changes to an open curve in the halfplane $(r \geq 0, z)$. Note that the function K_0 has the same logarithmic singularity as G_0, if $\vec{r} = (r,z) \to \vec{r}'$ [25, p. 67]. In the CVED-case, regularization is needed to reduce the problem of the disk scattering to the infinite-matrix equation of the Fredholm second-kind. This is done by applying a Galerkin scheme with the special Jacobi polynomials, which form the set of orthogonal eigenfunctions of IE static limit. In the CVMD-case, the basic IE is immediately of the Fredholm second kind provided that $R \neq 0$, and may be discretized via any reasonable projection scheme. In Figs. 10 to 12, the frequency scans of the power fractions related to the CVED-excited resistive flat-disk antenna on top of a grounded dielectric substrate are presented [26]. All the power values are normalized to

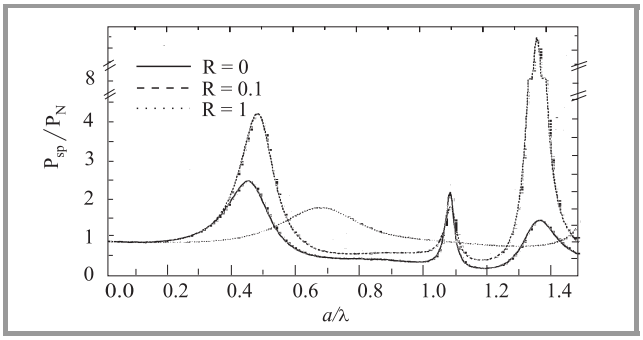


Fig. 10. Normalized radiated power as a function of the normalized frequency a/λ, for a VED-excited resistive circular-disk on a substrate. $\varepsilon_s = 1.07$, $h/a = 0.5$, and R as indicated.

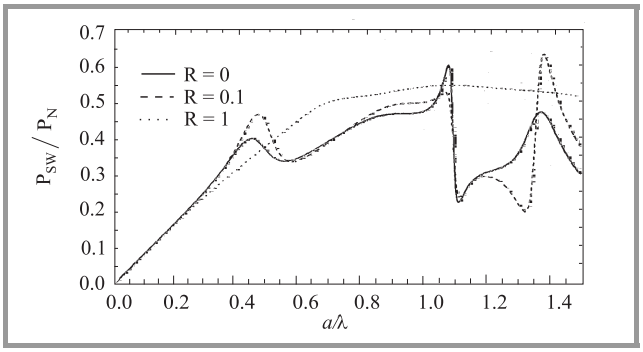


Fig. 11. Normalized surface-wave power as a function of the normalized frequency a/λ, for a VED-excited resistive circular-disk on a substrate. $\varepsilon_s = 1.07$, $h/a = 0.5$, and R as indicated.

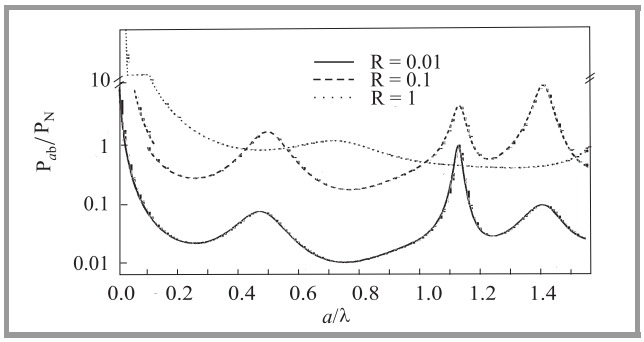


Fig. 12. Normalized absorbed power as a function of the normalized frequency a/λ, for a VED-excited resistive circular-disk on a substrate. $\varepsilon_s = 1.07$, $h/a = 0.5$, and R as indicated.

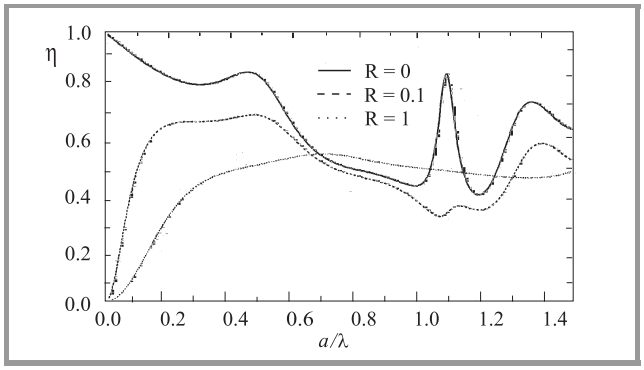


Fig. 13. Radiation efficiency of a VED-excited resistive disk antenna as a function of the normalized frequency.

the power radiated by a VED on a PEC plane. Figure 13 demonstrates the variation of the radiation efficiency in the same frequency band.

In the case of a thin flat material disk of high dielectric and magnetic constants, the boundary conditions (3) and (4) yield a set of two IEs, for the either type of coaxial excitation. Each of them is analogous to one of the resistive-disk IEs, therefore regularization and discretization is done as above. In [27], these IEs have been transformed to the

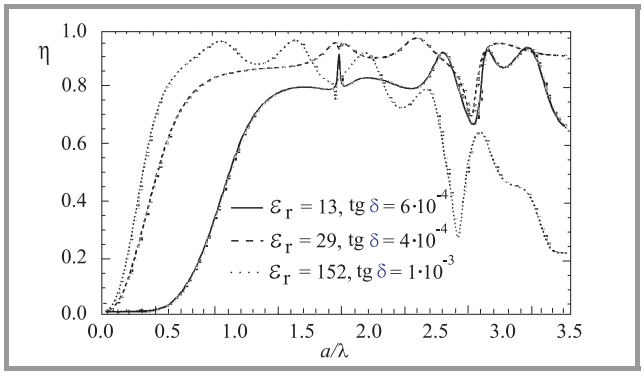


Fig. 14. Radiation efficiency of a thin-dielectric disk antenna as a function of the normalized frequency. $\varepsilon_s = 1.07$, $h/a = 0.05$, $b/a = 0.01$, $\mu_r = 1$, and ε_r as indicated.

Hankel-transform domain and converted to the dual integral equations for the surface-current transforms. Expansion functions are then transformed to the special-type Bessel functions depending on the type of IE. Numerical solution is very efficient and enables one to minimize the error to machine precision. Figure 14 shows the radiation efficiency of a dielectric disk antenna on a grounded dielectric substrate, fed by a CVED, as a function of the normalized frequency a/λ.

7. Conclusions

It is possible to modify the MAR solutions, previously developed in the PEC-screen scattering, to the imperfect thin screens: resistive, material, and impedance ones. This opens a way for a numerically exact analysis of not only the scattering but also the absorption of waves by the penetrable and impenetrable screens. We presented the illustrations related to the cases of curved strip, flat strip grating, and flat disk on a substrate. Some other results of analysis can be found in [15, 19, 22, 26, 27].

References

[1] Y. R. Grinberg, "Boundary conditions for electromagnetic field in the presence of thin metallic shells", *Radio Eng. Electron Phys.*, vol. 26, no. 12, 1981 (transl. Engl).

[2] G. Bouchitte and R. Petit, "On the concepts of a perfectly conducting material and of a perfectly conducting and infinitely thin screen", *Radio Sci.*, vol. 24, no. 1, pp. 13–26, 1989.

[3] G. Bouchitte, "Analyse limite de la diffraction d'ondes electromagnetiques par une structure mince", *C.R. Acad. Sci. Paris*, vol. 311, Ser. II, pp. 51–56, 1990.

[4] E. Bleszynski, M. Bleszynski, and T. Jaroszewicz, "Surface-integral equations for electromagnetic scattering from impenetrable and penetrable sheets", *IEEE Anten. Propag. Mag.*, vol. 36, no. 6, pp. 14–25, 1993.

[5] T. B. A. Senior and J. Volakis, *Approximate Boundary Conditions in Electromagnetics*. London: IEE Press, 1995.

[6] D. J. Hoppe and Y. Rahmat-Samii, *Impedance Boundary Conditions in Electromagnetics*. Taylor and Francis, 1995.

[7] T. B. A. Senior, "Diffraction by a semi-infinite metallic sheet", *Proc. Royal Soc. A*, vol. 213, pp. 436–458, 1952.

[8] R. F. Harrington and J. R. Mautz, "An impedance sheet approximation for thin dielectric shells", *IEEE Anten. Propag.*, vol. AP-23, no. 4, pp. 531–534, 1975.

[9] T. B. A. Senior, "Backscattering from resistive strips", *IEEE Trans. Anten. Propag.*, vol. AP-27, no. 6, pp. 808–813, 1979.

[10] R. C. Hall and R. Mittra, "Scattering from a periodic array of resistive strips", *IEEE Anten. Propag.*, vol. AP-33, no. 9, pp. 1009–1011, 1985.

[11] T. B. A. Senior and J. L. Volakis, "Sheet simulation of a thin dielectric layer", *Radio Sci.*, vol. 22, no. 7, pp. 1261–1272, 1987.

[12] T. B. A. Senior, "Generalized boundary and transition conditions and the question of uniqueness", *Radio Sci.*, vol. 27, no. 6, pp. 929–934, 1992.

[13] Z. T. Nazarchuk, *Singular Integral Equations in Scalar Diffraction Theory*. Lvov: Inst. Physics and Mechanics Press, 1994.

[14] A. I. Nosich, "Method of analytical regularization in wave scattering and eigenvalue problems: foundations and review of solutions", *IEEE Anten. Propag. Mag.*, vol. 42, no. 3, pp. 34–49, 1999.

[15] A. I. Nosich, Y. Okuno, and T. Shiraishi, "Scattering and absorption of E- and H-polarized plane waves by a circularly curved resistive strip", *Radio Sci.*, vol. 31, no. 6, pp. 1733–1742, 1996.

[16] E. I. Veliev, K. Kobayashi, T. Ikiz, and S. Koshikawa, "Analytical-numerical approach for the solution of the diffraction by a resistive strip", in *Proc. Int. Symp. Anten. Propag. (ISAP-96)*, Chiba, 1996, pp. 17–19.

[17] M. M. Lavrentyev, *Some Ill-Posed Problems of Mathematical Physics*. Novosibirsk: SO AN SSSR Publ., 1962 (in Russian).

[18] A. B. Bakushinsky, "About one numerical method of solving the Fredholm first-kind integral equations", *USSR J. Comput. Mat. Math. Phys.*, vol. 5, no. 4, pp. 744–749, 1965.

[19] A. I. Nosich, V. B. Yurchenko, and A. Altintaş, "Numerically exact analysis of a 2-D variable-resistivity reflector fed by a complex-point source", *IEEE Trans. Anten. Propag.*, vol. AP-45, no. 11, pp. 1592–1601, 1997.

[20] E. I. Veliev, S. Koshikawa, and K. Kobayashi, "Diffraction of a plane wave by a thin material strip: solution by the analytical-numerical approach", in *Proc. Int. Conf. Math. Meth. EM Theory (MMET'2000)*, Kharkov, 2000, vol. 1, pp. 189–192.

[21] R. Petit and G. Tayeb, "Theoretical and numerical study of gratings consisting of periodic arrays of thin and lossy strips", *J. Opt. Soc. Am. A*, vol. 7, no. 9, pp. 1686–1692, 1990.

[22] T. L. Zinenko, A. I. Nosich, and Y. Okuno, "Plane wave scattering and absorption by resistive-strip and dielectric-strip periodic gratings", *IEEE Trans. Anten. Propag.*, vol. AP-46, no. 10, pp. 1498–1505, 1998.

[23] A. Matsushima, T. L. Zinenko, H. Minami, and Y. Okuno, "Integral equation analysis of plane wave scattering from multilayered resistive strip gratings", *J. Electromagn. Wav. Appl.*, vol. 12, no. 10, pp. 1449–1469, 1998.

[24] S. L. Prosvirnin, S. A. Masalov, A. V. Ryzhak, and V. M. Shkil, "Electromagnetic wave diffraction by a flat grating of resistive strips", *Radiotekhn. Elektron.*, vol. 43, no. 7, pp. 792–796, 1998 (in Russian).

[25] V. I. Dmitriev and E. V. Zakharov, *Integral Equations in Electromagnetic Problems*. Moscow: Moscow State University Press, 1987 (in Russian).

[26] N. Y. Bliznyuk and A. I. Nosich, "Numerical analysis of a lossy circular microstrip antenna", in *Radio Physics and Electronics*. Kharkov: IRE Press, 1999, vol. 4, no. 3, pp. 125–128.

[27] N. Y. Bliznyuk and A. I. Nosich, "Numerical analysis of a dielectric disk antenna", in *Radio Physics and Electronics*. Kharkov: IRE Press, 2000, vol. 5, no. 1, pp. 49–54 (in Russian).

Alexander I. Nosich was born in 1953 in Kharkov, Ukraine. He had graduated from the Kharkov State University and earned his M.Sc., Ph.D. and D.Sc. (higher doctorate) degrees, all in radio-physics, in 1975, 1979 and 1990, respectively. Since 1979, he is on research staff of IRE NASU as a scientist. In 1992–2000, he held visiting professorships and guest-scientist fellowships in Turkey, Japan, Italy and France. In 1995, he was the organizer and first chairman of the IEEE AP Society East Ukraine Chapter (now joint with MTT, ED, AES, GRS and NPS Societies), the first one in the Former Soviet Union. In 1996–2000, he was the chapter secretary-treasurer. Since 1995, he is on the editorial board of the *Microwave and Optical Technology Letters*. In 1990–2000, he was an organizer and technical program committee co-chairman of the series of international conferences on *Mathematical Methods in Electromagnetic Theory (MMET)* in the USSR and Ukraine. His research interests include wave scattering, radiation, propagation and absorption studied by the analytical regularization techniques, and the history of microwaves.
e-mail: alex@emt.kharkov.ua
Institute of Radiophysics and Electronics
National Academy of Sciences
Ulitsa Proskury 12, Kharkov 61085, Ukraine

Regular paper

Optimum double-matched detection and its application to SSMA systems

Józef Jacek Pawelec

Abstract — This paper is concerned with the issue of optimum detection of known signal in nonwhite noise and/or narrow-band interference. The detection is carried out in three steps. First, some function related to the power spectrum of interfering process *I(z)* is estimated via adaptation. Second, the signal + interference is whitened due to *I(z)*. Third, the replica of signal is filtered via *I(z)* to match it to the deformed signal in the previous step. The simulation of SS reception in presence of NB interference shows a high gain in comparison to classical single-matched detection.

Keywords — optimum detection, spread spectrum multiple access systems, NB interference, adaptation.

1. Introduction

The issue of optimum detection is nearly as old as the telecommunication itself [1]. There is a lot of methods and criteria of optimality, e.g. integration or storage, autocorrelation, coherent or synchronous detection, optimum linear or matched filter detection, statistical detection via risk strategies (Bayes, Neyman-Pearson and mini-max) [2 – 4]. Of course, there are some common roots in all methods.

We will confine ourselves to the matched-filter and maximum-likelihood (ML) strategies. In contemporary radio receivers the matched-filter philosophy is widely used but is confined to the useful signal only. It is a reasonable solution for some kind of communications corrupted by AWG noise. In wireless communication, especially spread spectrum (SS) the dominant factor is an outside noise and it diverges considerably from the thermal white model. The resulting miss-adjustment between "white" receiver filter and nonwhite noise/interference causes losses that reach even several tens of dB [5, 6].

This problem has been considered in [7 – 10]. In the first paper [7] an optimum structure of SS receiver is analyzed. Channel distortion and multi-point reception are studied here in detail, whilst colored interference is merely touched on (a simple low-pass case is analyzed). A substantial, but extremely concise approach to nonwhite detection is given in [8]. It contains, however, no functioning structures, no algorithms, and nor comparative results. This gap is partially fulfilled by the author's studies [9] and [10], which the present contribution is based upon.

The organization of the paper is as follows. In Section 2, the statement of the problem in modern digital form is given. Thereafter a new general detection structure is introduced. In Section 3 and in Appendix A, the adaptive theory, relevant to the matter is provided and the general scheme is complemented and simplified. In Section 4, the simulation of spread spectrum reception in narrow-band interference is carried out and the results obtained are compared with conventional reception.

2. Statement of the problem

As it was shown by many authors [3, 8, 10], the optimum filter for detection of a given signal $S(\omega)$ in presence of arbitrary noise is

$$\mathsf{H}(\omega) = \frac{S^*(\omega)}{P(\omega)}, \tag{1}$$

where $S^*(\omega)$ is a conjugate spectrum of useful signal and $P(\omega)$ – power spectrum of interfering process.

The denominator of the desired transfer function, $P(\omega)$ is usually factorized as $P(\omega) = H(\omega)H^*(\omega) = |H(\omega)|^2$ or – within the z variable – as [8]

$$P(z) = A^2 G(z) G^*(1/z^*), \tag{2}$$

where $AG(z)$ corresponds to $H(\omega)$, $AG^*(1/z^*)$ – to $H^*(\omega)$; A is a constant; $z = re^{j\Omega}$; $-\pi \le \Omega \le \pi$, $r > 0$ (observe the difference between $\mathsf{H}(\omega)$ and $H(\omega)$).

$G(z)$ is said to be a minimum-phase function, while $G^*(1/z^*)$ – maximum-phase function. As $G(z)$ is minimum-phase, all its zeros are inside the unit circle and – after inversion – they do not reach the unstable region outside this circle. Hence, the reciprocal of $AG(z)$ can be easily specified as $I(z) = 1/AG(z)$ and put in series with the signal matched-filter $S_l^*(1/z^*)$, see Fig. 1 (l is an alphabet index).

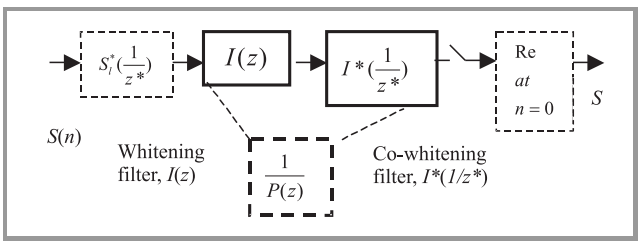


Fig. 1. A general idea of double-matched detection [7].

The $G^*(1/z^*)$ is – by definition – maximum-phase and has no reciprocal. However, its inverse Fourier transform h^*_{-k} (said anti-causal response) can be shifted some N steps forward without affecting the transform (k is a discrete time index). This way a causal response h^*_{-k+N} is obtained, which – after reversion – can play a role of replica in cross-correlation mode (Fig. 2). Consequently, the co-whitening filter in matched-filter mode stands for a whitening filter in cross-correlation mode, $I_{cowhite}(z) \to I_{white}(z) = I(z)$ and $h^*_{-k+N} \to h_k$ (Fig. 2).

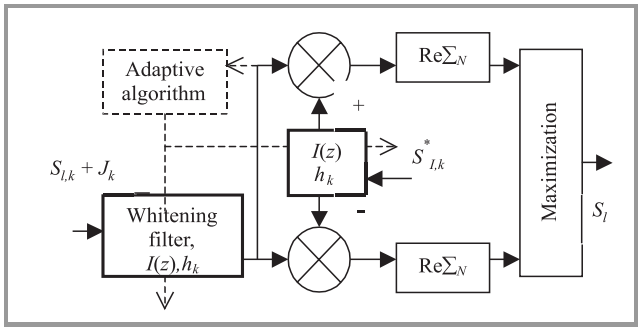


Fig. 2. A cross-correlation mode of scheme Fig. 1 for binary signal and unknown interference [8].

An example. Let $P(z) = 1/|1 + 0.5\ z^{-1}|^2$. Hence, the whitening and co-whitening functions are: $I(z) = 1 + 0.5\, z^{-1}$ and $I^*(1/z^*) = 1 + 0.5\, z^{+1} (A = 1)$ [8]. Let an input signal be $\mathsf{s}(n) = [1\ 1\ 1\ \ldots\ 1]$ for time sequence $n = [-5, 5]$ otherwise $\mathsf{s}(n) = 0$. In this case the output signal of $S^*_l(1/z^*)$ will be the same as its input, since we assume $h_k = 1$. The transition of $s(n)$ through $I(z)$ yields $s'(n) = [1\ 1.5\ 1.5\ \ldots\ 0.5]$ for $k = [-5, 6]$. The $I^*(1/z^*)$ is maximum-phase, so we factorize it in two functions, $I_1(z) = z^{+1}$ and $I_2(z) = 0.5 + z^{-1}$. Of course, $I(z) = I_1(z) I_2(z)$. The first function $I_1(z)$ shifts $s'(n)$ one step ahead, i.e. $s''(n) = [1\ 1.5\ 1.5\ \ldots\ 0.5]$ for $n = [-6, 5]$. The second function transforms $s''(n)$ into $s'''(n) = [0.5\ 1.75\ 2.25\ \ldots\ 2.25\ \mathbf{2.25}\ 2.25\ \ldots\ 2.25\ 1.75\ 0.5]$ for $n = [-6, 6]$. The decision is taken at $n = 0$, hence the result is $S = 2.25$.

Similar operations in cross-correlation mode (Fig. 2), yield $S_\tau = \sum s'(n) s^{*'}(n) \Delta\tau$, where $s^{*'}(n) = s'(n)$ and $\Delta\tau$ is a time interval between steps. Putting $\Delta\tau = 1/10$ we obtain $S_\tau = 2.15$. For a more dense digitalization $(N >> 10)$ $S_\tau \approx S$. We will not further deal with the errors appearing in this process, as the signals feeding the modern detectors are digital in nature at the very origin. The same decrements refer to the useful signal as refer to the noise and SNR is held constant.

The scheme Fig. 2 works as follows. An lth useful signal $S_{l,k}$ and the interference J_k enter the whitening filter $I(z)$. It decorrelates J_k and changes $S_{l,k}$. The first action is desirable, as it enables the application of ML principle. The second – is undesirable, so the replica $S^*_{l,k}$ is passed through the same filter $I(z)$ to match it to the deformed signal $S_{l,k}$. The output products of filters, y_k and y^*_k are multiplied one by one and summed up within the range of $k = 1, 2, \ldots, N$. This process is repeated for the full alphabet of signals, $l = 1, 2, 3, \ldots$. The real values of sums are then compared each other and this l is taken as true one, which assigned sum is maximal. At the scheme Fig. 2 we have shown only the simplest binary case for $l = \pm 1$. Now, let us consider the constraints that have been implicitly imposed upon the considered processes.

- First, the power spectrum of interference $P(z)$ has to be determined and/or the whitening stable function $I(z)$ has to exist (even, in approximated form).
- Second, the input signal $S_{l,k}$ and the replica $S^*_{l,k}$ have to be strictly synchronized one to another and not correlated to the interference J_k.
- Third, the interference J_k has to be wide sense stationary (WSS) or at least cycle-stationary (within the adaptive period) and not necessarily Gaussian, whilst possibly nonwhite. The Gaussianity is unnecessary as the interference goes through filter $I(z)$, which normalizes it.

All the above requirements are usually satisfied, except for the first. The power spectrum of interference $P(z)$ and the whitening function $I(z)$ are often unknown, so we will try to get them in a course of adaptation.

3. Blind adaptation

The classical adaptive procedure consists in a comparison of a received signal with some standard, e.g. a training sequence. This comparison produces an error, which controls the weights in adaptive filter. In our position, no standard signal exists, so only the blind adaptive algorithms can be considered.

The simple blind algorithm, belonging to the large minimum mean square error group (MMSE) is as follow [10]. At the begining, an estimate of the interference sample $\hat{y}(n)$ is formed upon its previous states (observations) $y(n-N)$, $y(n-N+1)$, ..., up to $y(n-1)$, Eq. (3). In the next step, this estimate is subtracted from the actual value of $y(n)$ (both are accessible), Eq. (4). The difference $\varepsilon(n)$ is used for the step by step matching of filter weights $\boldsymbol{h}$ to observed process, Eq. (5). The algorithm is sometimes called least squares (LS) [7, 12]

$$\hat{y}(n) = \sum_{k=1}^{N} h_k y(n-k) = \mathbf{h^T y}\,,$$

$$\mathbf{y} = \begin{bmatrix} y(n-1) \\ \cdots\cdots\cdots \\ y(n-N) \end{bmatrix}, \quad \mathbf{h} = [\mathsf{h}_1\ \mathsf{h}_2\ \ldots\ \mathsf{h}_k\ \ldots\ \mathsf{h}_N]^{\mathsf{T}}, \quad (3)$$

$$\varepsilon^2(n) = Ex\left\{ \left[y(n) - \hat{y}(n) \right]^2 \right\}, \quad (4)$$

$$\mathbf{h}(n+1) = \mathbf{h}(n) - \mu\, \varepsilon \mathbf{y}\,. \quad (5)$$

The step size μ, as well as an initialize sequence $\boldsymbol{h}(0)$ are important data, which affects the convergence and the speed of adaptative process. We experienced to use $\boldsymbol{h}(0) = [1\,0\,1\,0\,\ldots]^T$ and $\mu = 0.01 \sim 0.001$. The higher values of μ correspond to the higher speeds of adaptation and the smaller – to smaller errors of estimation. In any case $\mu_{\max} < 2/\lambda_{\max}$, where λ is an eigenvalue of the autocorrelation process.

The estimation time of MMSE algorithms is usually large and the orders of filter – very high. We used the sequences of size $N = 10^3 \sim 3 \cdot 10^3$ and $L = 6 \sim 46$. The higher values of L and N correspond to narrow pulses of interference. For the transmission rates of $10 \sim 30$ kbit/s it stands for an estimation time $\tau = 0.1$ s. More effective algorithms are nonlinear ones, e.g. gradient descent, super exponential and turbo [15]. They use – instead of square error – a notion of cost function. An example of such function is a normalized moment of fourth order:

$$u(4) = \frac{m_4(y)}{m_2^2(y)} = (N-L)\,\frac{\sum\big[y(n)-Y\big]^4}{\Big\{\sum\big[y(n)-Y\big]^2\Big\}^2}, \qquad (6)$$

where $y(n)$ is time series of estimated signal; N – series size; L – order of whitening filter.

$$Y = \frac{1}{N-L}\sum_{n=L+1}^{N} y(n)\,, \qquad (7)$$

$$m_2(y) = \frac{1}{N-L}\sum_{n=L+1}^{N}\big[y(n)-Y\big]^2, \qquad (8)$$

$$m_4(y) = \frac{1}{N-L}\sum_{n=L+1}^{N}\big[y(n)-Y\big]^4. \qquad (9)$$

The moment $u(4)$ for commonly used bipolar signal, $y \in [1, -1]$ takes the lowest value equal 1. In a physical channel, the signal samples interact with one another, which causes cross-correlation (ISI) and gives an increase of $u(4)$. So, if we want to decorrelate (white) the signal, a gradient of $u(4)$ has to be used as an indicator pointing the direction, to which the vector $\boldsymbol{h}$ has to be changed.

A weak point of some nonlinear algorithms is their false convergence. If "the spectral channel" of interference is non minimum-phase and reveals several local minima, one of them (not global minimum) can cause a false convergence. In Appendix A, we present the very effective lattice filter and the gradient descent algorithm (GDA) that do not reveal a false convergence.

4. Simulation experiments

We are carrying out several experiments to have a deeper insight into the filtration and detection processes and to estimate a gain of the method. We used the spread spectrum signal of the form

$$S_{l,k} = l \times \big[s_1, s_2, \ldots, s_N\big]_1, \ldots, l \times \times \big[s_k, s_{k+1}, \ldots, s_{k+N-1}\big]_{(k\,mod\,N+1)}, \ldots, l \times \times \big[s_{MN-N+1}, \ldots, s_{MN}\big]_M, \qquad (10)$$

where l is a binary random number, $l \in [1, -1]$, its sequence $\{l\}$ contains the information carried on; N is a spreading factor; $[s_k, s_{k+1}, \ldots]_{k\,mod\,N+1}$ represents mth bit of signal, $m = 1, 2, 3, \ldots, M$; s_k is kth sample of signal, $k = 1, 2, \ldots, MN$; $s_k \in [1, -1]/\gamma$, γ – real constant; the sequence $[s_k]$ is chosen according to Gold code and is known to the sender and recipient of information.

The appropriate interference sequence is $J_k = \{J_1, \ldots, J_{MN}\}$, where MN is a sequence size in chips. The replica $S^*_{l,k}$ follows the useful signal except for the attributes of power (γ), information carried on (l), and conjugation index $(^*)$, $|s^*_{1,k}| = |\gamma s_{l,k}|$.

We have carried out four experiments. The first one is concern with the white noise case. The symbol of sum in Fig. 2 denotes

$$\sum_N = \sum_{k=1+jN}^{N+jN} y_k y_k^*, \quad j = 0, 1, 2, \ldots, M-1\,, \qquad (11)$$

where y_k and y_k^* are the outputs of whitening filters in the signal and replica leads, respectively.

The obtained BER curve is shown in Fig. 3 as the base curve (0). We found that interference other than white, e.g. low-pass of $H(z) = 1/(1 - 1.4\,z^{-1} + 0.5\,z^{-2})$ causes a degradation of reception, expressed by curve (1). This is contrary to Shannon's theory, which states that white noise causes the highest degradation of reception. So, we insert a proper whitening filter into both paths of detector, $I(z) = 1/H(z) = 1 - 1.4\,z^{-1} + 0.5\,z^{-2}$ (experiment II). The symbol of sum denotes now

$$\sum_N = \sum_{k=1+L+jN}^{N+L+jN} y_k y_k^*, \qquad (12)$$

where L – order of filter $(= 2)$.

The obtained BER curve is shown as curve (2), see Fig. 3. We observe that it has been shifted about 7 dB down the base curve (0). This is a striking difference and it means that the single-matched (white) detector does not recognize the color of interference.

Then, further experiments were carried out. In the third one, III, an interference pulse of natural signal was used. Let its autocorrelation function and power spectrum be (e.g. BPSK signal)

$$R(\tau) = 1 - |\tau|/T \text{ for } |\tau|/T < 1, \text{otherwise } R(\tau) = 0\,, \quad (13a)$$

$$P(\omega) = T\big[\sin(\omega T/2)/(\omega T/2)\big]^2. \qquad (13b)$$

Having $P(\omega)$ we can determine its corresponding transfer function $H(z)$ and the unit impulse response $\boldsymbol{h}$ via Yule-Walker method [18]

$$[\mathbf{a}, \mathbf{b}] = yulewalk\,(L, \mathbf{F}, \mathbf{M}),\ \mathbf{h} = impz\,(\mathbf{a}, \mathbf{b})\,, \qquad (14)$$

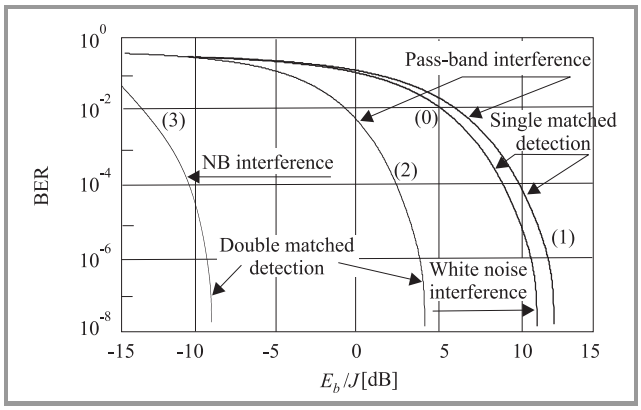


Fig. 3. A family of BER curves for single-matched and double-matched detection at different interference.

where L is a chosen order of filter; $\mathbf{F}$ – frequency scale vector, e.g. $\mathbf{F} = [0\ 0.01\ \ldots\ 1]^T$; $\mathbf{M}$ – interference magnitude vector, $\mathbf{M} = |H(\mathbf{F})|^T$; $\boldsymbol{a}$, $\boldsymbol{b}$ – transfer function coefficients vectors.

The formulae for $H(z)$ and its coefficients are

$$H(z) = \frac{a_0 + a_1 z^{-1} + \ldots + a_{16} z^{-16}}{1 + b_1 z^{-1} + \ldots + b_{16} z^{-16}}, \tag{15}$$

$\boldsymbol{a} = [0.32\ \ -0.01\ 1.35\ \ -0.04\ 2.7\ 0.07\ 3.36\ \ -0.07\ 2.7\ \ -0.05\ 1.48\ \ -0.02\ 0.5\ \ -0.005\ 0.1\ \ \sim 0\ \ \sim 0]^T$,

$\boldsymbol{b} = [1\ \ -0.037\ 5.5\ \ -0.18\ 13.8\ \ -0.39\ 20.4\ \ -0.49\ 19.5\ \ -0.38\ 12.3\ \ -0.18\ 5.03\ \ -0.05\ 1.2\ \ -0.006\ 0.13]^T$.

To obtain $I(z)$ we simply put $\boldsymbol{b}$ instead of $\boldsymbol{a}$ and *vice versa* into a filter $(\cdot)$ function [18]. The result we obtained in experiment III is expressed by curve (3), see Fig. 3. This curve has been shifted about 20 dB down the base (0). Although hard to imagine, it is true. The gain obtained can be expressed by an approximated formula

$$g \approx 20\ \log_{10} B/B_i, \tag{16}$$

where B is useful signal equivalent bandwidth and B_i – interference equivalent bandwidth.

In the next experiment (IV) we assumed that interference spectrum is not known, so, the adaptive loop was activated (dotted line block, Fig. 2). In a course of MMSE adaptation we obtained $\hat{\boldsymbol{h}}$ very similar to $\boldsymbol{h}$ of the previous experiment (a difference in E_b/N_0 was less than 1 dB). This happened because we used the same interference pulse both in experiments III and IV, simply to check the adaptive process.

Many other interference pulses (QPSK, QAM) with different bandwidths and locations on signal spectrum (including double pulses) were examined (Table 1). All results were very good [8]. Several experiments have been carried out using the lattice filters and the gradient descent algorithm (Appendix A, Fig. 4). We found that in this case the order of filter, and the time of adaptation diametrically reduce, and the problem of instability disappears. However, the program is more complicated, and the physical implementation of lattice filters would possibly be more complex.

Table 1
Chosen SNR-s for presence and absence of whitening filters and the different locations of interference on the signal spectrum ($B/B_i = 10$)

	E_b/J [dB] at BER = 10^{-5}				
Interference pulse relative location	0.0	0.25	0.5	0.75	1
No whitening filters	10	10	10	10	10
Filters present	−7.5	−8.5	−9.5	−8.5	−7.5

It should be also noted that a described process is strictly optimal only for known interference spectrum. Otherwise, when using adaptation, the algorithm "learns" the spectrum and if its "equivalent channel" is non minimum-phase and the whitening filter is FIR, then the detector can only be referred to as asymptotically optimum (for $L \to \infty$) [12]. In spite of so rigorous theoretical limit, the practical requirements are not so excessive, e.g. $L = 46$ is enough for a relative signal-to-interference bandwidth $B/B_i = 10$.

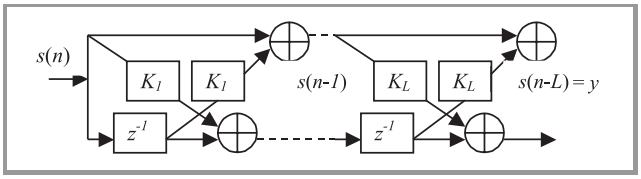


Fig. 4. A structure of whitening lattice filter (see Appendix A).

The problem of optimum adaptive algorithm and adaptive filter is still an open issue. But, even for slow MMSE algorithm we can obtain quite satisfactory results under typical communication conditions.

5. Conclusion

This paper presents a new method of spread spectrum detection in presence of nonwhite noise and/or narrow-band interference. It uses a double-matched detector, fitted both to the useful signal and interference. It is formed by adding a pair of whitening filters and a blind adaptive loop to the conventional single-matched ML structure. Simulation of base-band SS reception under typical NB interference yields the gain approaching 20 dB for the signal-to-interference bandwidth ratio $B/B_i \approx 10$. Unfortunately, the gain reduces to zero for flat interference.

The obtained results are comparable to the findings of other authors [5, 6]. The method is intended for systems with interference dominating inner thermal noise and being ruggedly nonwhite. It can be treat as a compliment to RAKE and other upgrading techniques.

Appendix A
Gradient descent algorithm and the lattice filter

The normalized moment $u(4)$ and its derivatives are

$$u(4) = \frac{m_4(y)}{m_2^2(y)},\ \frac{du}{d\mathbf{K}} = \frac{du}{dy}\frac{dy}{d\mathbf{K}},\ \mathbf{K} = [K_1 K_2 \ldots K_L]^{\mathrm{T}}, \quad (1A)$$

were $\boldsymbol{K}$ is so-called reflection coefficient in lattice filter, see Fig. 4 (it corresponds to the weight vector $\boldsymbol{h}$ in transversal FIR filter); y is an output signal of filter.

Taking into account Eqs. (12) and (13) in main text, we obtain (2A)

$$\frac{du}{dy} = \frac{m_4' m_2^2 - 2m_2 m_2' m_4}{m_2^4} =$$
$$= \frac{4}{(N-L+1)m_2^2}\sum\left[y(n)-Y\right]^3 +$$
$$- \frac{4m_4}{(N-L+1)m_2^3}\sum\left[y(n)-Y\right]\Big\} = U - V. \quad (2A)$$

As it comes from Fig. 4 (derivation in [11])

$$\frac{dy}{dK_1} = s(n-1)(1+K_2) +$$
$$+ s(n-2)(0+K_3) + s(n-3)(0+\ldots, \quad (3A)$$

$$\frac{dy}{dK_2} = s(n-1)(K_1+K_3) +$$
$$+ s(n-2)(1+K_4) + s(n-3)(0+\ldots, \quad (4A)$$

$$\frac{du}{dK_p} = \left\{U-V\right\}\left\{s(n-1)\left[K_{p-1}+K_{p+1}\right] +\right.$$
$$+ s(n-2)\left[K_{p-2}+K_{p+2}\right] + \ldots\Big\} =$$
$$= \left\{U-V\right\}\sum_{i=1}^{L} s(n-i)\left[K_{p-i}+K_{p+i}\right],$$

where $s(\cdot)$ is input signal of the filter, and

$$K_0 = 1,\ K_{p\pm i} = 0 \ \text{ for } \ p \pm i < 0 \ \text{ or } \ p \pm i > L. \quad (5A)$$

The quantities U, V are given in (2A). The fundamental gradient equation is

$$\mathbf{K}(n+1) = \mathbf{K}(n) - \mu\nabla_K u(4), \quad (6A)$$

$$\nabla_K u(4) = \left[\frac{du(4)}{dK_1}\ \frac{du(4)}{dK_2}\ \ldots\right]^T, \quad (7A)$$

where μ is a step size and n – iteration number. The convergence criteria, initialize parameters and a step size μ are considered in [12, 15].

Acknowledgment

The author is grateful to the unanimous Reviewer for his insightful comments and helpful critiques of the manuscript. This work was partially supported by a grant of State Committee on Science Research, no. 0T00A-037-16/98.

References

[1] R. Courant and D. Hilbert, *Methoden der Mathematischen Physik*. Berlin: Springer, 1931.

[2] W. Davenport and W. Root, *An Introduction to the Theory of Random Signals and Noise*. McGraw-Hill, 1958.

[3] C. Helstrom, *Statistical Theory of Signal Detection*. Pergamon Press, 1960.

[4] P. Beckman, *Probability in Communication Engineering*. Harcourt, 1967.

[5] V. Comley, "CW interference excision in DSSS communication system using spectrally defined spreading/despreading functions", in *IEEE Milit. Commun. Conf. MILCOM'98*, Bedford, Oct. 1998, vol. I, pp. 160–164.

[6] R. Derryberry, T. Wong, and J. Lehnert, "An iterative blind adaptive receiver for DS-SSMA systems", in *IEEE Milit. Commun. Conf. MILCOM'98*, Bedford, Oct. 1998, vol. II, p. 499–503.

[7] A. Reichman and R. Scholtz, "Adaptive spread-spectrum systems using least-squares lattice algorithm", *IEEE J. SAC*, vol. 3, no. 5, pp. 653–662, 1985.

[8] E. Lee and D. Messerschmitt, *Digital Communication*. Boston: Kluwer, 1997.

[9] J. Pawelec, "Optimum adaptive detection of SS signals in nonwhite noise/interference", in *Int. Conf. Comput. Electromagn. Its Appl.*, Beijing, Nov. 1999, pp. 522–524.

[10] J. Pawelec, "Optimum adaptive detection of SS signals in NB interference", *Bull. Defen. Acad. Technol.*, no. 3, 2000 (also in *ISSSE Symp.*, Tokyo, July 2001).

[11] J. Pawelec and A. Janulewicz, "Non-Gaussian signals separation via HOS", in *Int. Symp. EMC*, Zurich, Feb. 1999.

[12] S. Haykin, Ed., *Adaptive Filter Theory*. Prentice Hall, 1996.

[13] O. Shalvi and E. Weinstein, "New criteria for deconvolution of non-minimum phase systems", *IEEE Trans. Inform. Theory*, vol. 36, no. 2, pp. 312–321, 1990.

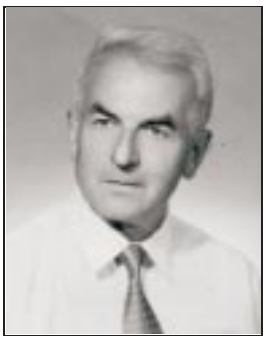
[14] J. Cadzow, "Blind deconvolution via cumulant extrema", *IEEE Sig. Proc. Mag.*, May 1996.

[15] C. Jonson *et al.* Special issue on blind systems. *Proc. IEEE*, vol. 86, no. 10, 1998.

[16] J. Holmes, *Coherent Spread Spectrum Systems*. Wiley, 1982.

[17] B. Sklar, *Digital Communications*. Prentice Hall, 1988.

[18] Matlab 5.1/0421,Toolbox: Signal Proc. The Mathematical Works Inc. 1997.



Józef Jacek Pawelec received the M.Sc. degree in 1958 in radio engineering from Military Academy of Technology, Warsaw. The Ph.D. and D.Sc. degrees received in 1975 and 1982, respectively from the Academy, Faculty of Electronics. In 1982–90 he was a deputy director of Defence Communication Institute and in 1990–94

the chair of electronic section at Academy. Now he is a Professor at University of Technology Radom. Research interests: electromagnetic compatibility, radio communication systems, signal processing (detection, blind adaptation). A book: "Control and communication in space", WKŁ. More than 100 scientific papers (a part published abroad). A member of National Committee of URSI.
e-mail: pawelec@wil.waw.pl
Defence Communication Institute
05-130 Zegrze, Poland

Regular paper

Photo-devices for optical controlling of microwave circuits

Zenon R. Szczepaniak and Bogdan A. Galwas

Abstract — The most important optical devices which can be used for controlling microwave circuits will be presented in the paper. The performance and the parameters of the devices such as semiconductor microwave optoelectronic switches, photodiodes and phototransistors were described. The influence of the optical illumination on their microwave parameters will be described in details, including the our own investigations and simulations results. Several applications of such devices and their potential possibilities will be presented.

Keywords — microwave optoelectronic switch, photodiode, phototransistor.

1. Introduction

During last years it can be seen a new trend in the telecommunication field – merging the optical and the microwave techniques. It means that there is a need to drive the optical devices with the high frequency or microwave signals. On the other hand we need to control the microwave circuits by means of the optical signals. Such optically controlled microwave circuits are wanted for optically provided processing of microwave signals. They are strongly needed for new types of radar. Also the optically controlled microwave mixers and oscillators are used in hybrid-fibre radio applications [1].

One of the solutions for optical control of microwave circuits is the developing of the microwave devices with the capability of changing their microwave parameters such as capacitance, impedance, reflectance or current source value due to the absorption of the light. Microwave circuit designer usually can propose a circuit, with the use of such a device, which will have the possibility of changing its attenuation/gain, or generation frequency by means of the illuminating optical power. The various families of such devices will be described in the following sections of the paper.

2. Bulk photoconductive devices

2.1. Photoconductive effect

The most common application of the illuminated bulk semiconductor devices is switching. The idea is to use a piece of semiconductor material to make a gap in a microwave transmission line, and then the switch is activated by illumination. The principle of the operation relies on the light absorption in the semiconductor material. When the incident photon energy is higher than the band gap of the semiconductor, the photon is absorbed and the creation of electron-hole pair takes place. This phenomenon causes the increase of the carrier density in the semiconductor material under the optical illumination, and therefore the increase of the semiconductor conductivity. Light absorption coefficient α depends on the type of semiconductor and also on the light wavelength λ (Fig. 1):

$$\alpha\ [\mathrm{cm}^{-1}] = 2 \cdot 10^4 \sqrt{hf - E_g\ [\mathrm{eV}]}, \qquad (1)$$

where h is the Planck's constant, f is a frequency of the incident light and E_g is a band gap energy of semiconductor material.

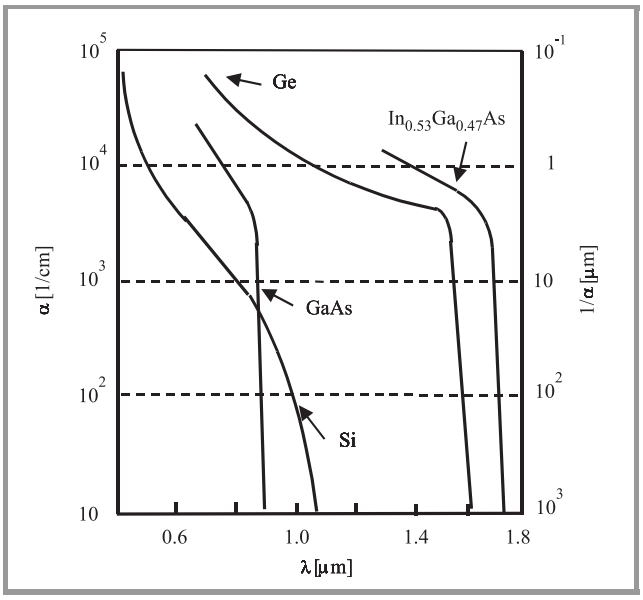


Fig. 1. Light absorption coefficient α and penetration depth $1/\alpha$ versus wavelength λ for different types of semiconductor material.

The absorption of an optical power P_0 in the semiconductor material can be described by the exponential law, as it is expressed by relation (2):

$$P_{abs} = P_0(1-\rho)\,(1-e^{-\alpha x})\,. \qquad (2)$$

It is important to notice that the optical power penetrating the semiconductor material is lower than P_0 because some part ρ of the optical power is reflected from the semiconductor surface due to light refraction coefficient difference.

2.2. Photoconductive microwave switches

The basic construction and idea of the operation of the photoconductive microwave switch is shown in Fig. 2. The PC switches base on the transmission line, coplanar line or

microstrip one, fabricated on semiconductor [2 – 4]. The conducting strip of the coplanar line has a gap causing the semiconductor to be exposed to the optical illumination.

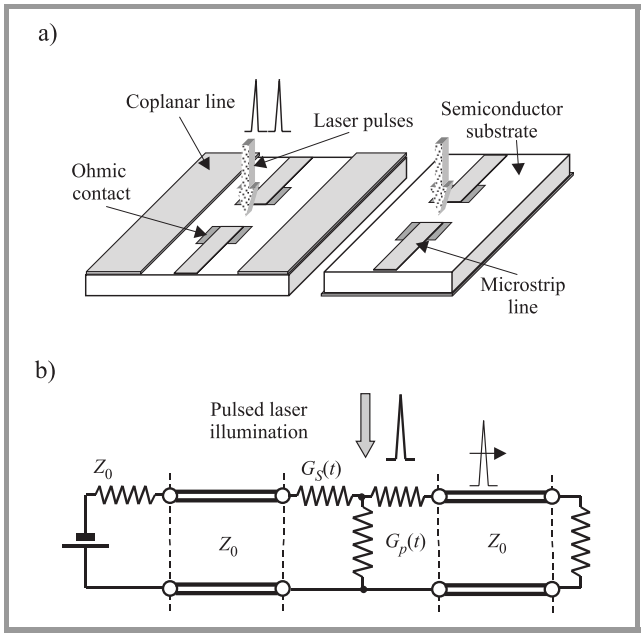


Fig. 2. Schematic illustration of a photoconductive microwave switch (a), and equivalent circuit (b).

When the light illuminates the gap area, such that its photon energy is greater than the semiconductor band gap, the light is absorbed and the free electron-hole pairs are generated. Thus, the conductivity of the semiconductor is altered. The illumination creates the surface plasma in the semiconductor. This surface plasma enhances the surface conductivity and bridges the gap in the strip conductor. Then the microwave signal can propagate through the switch.
In Fig. 2b the equivalent circuit of PC switch is shown. The values of the conductances series G_S and parallel G_P are the function of the optical power and are modulated by laser pulses.

Example: The typical parameters of a microstrip switch made on a silicon substrate are: refraction coefficient of semiconductor material $n_r = 3.6$, electron and holes mobilities $\mu_n + \mu_p = 2000$ cm$^2 \cdot$V$^{-1} \cdot$s^{-1}, strip metalisation gap $L = 0.034$ cm, switching ON pulse energy $E_1 = 10$ μJ and $hf/e = 2.34$ V. Then the optically generated conductivity in the microstrip gap is equal to 5 S (it means 0.2 Ω). Thus in the ON state the optoelectronic microstrip switch has microwave reflection coefficient with magnitude equal to -54 dB and in the OFF state equal to -0.06 dB. They are good results.
The photoconductive microwave switches promise faster rise and fall time (typically in the picosecond range), broader bandwidth, ability to handle high power, and simplicity of operation in comparison to the conventional GaAs MESFET or PIN diode switches. Another advantage is the electrical isolation between the controlling optical signal and the gated microwave signal.

The very short electrical pulses generated with the use of the PC switches are used for measurements in the millimetre and sub-millimetre bands, and for the time domain measurements of the dielectric materials (Fig. 3), antennas, transistors and microwave amplifiers, especially if they work far from bands of the frequency domain network analysers.

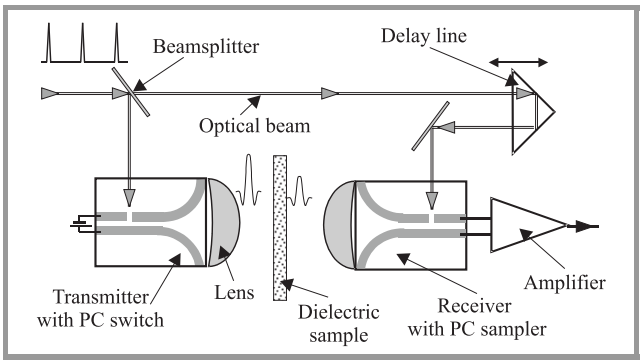


Fig. 3. Configuration of the system for transmission measurements of the dielectric properties of materials in the millimetre-wave region.

There was some investigations of the use of the variable impedance provided by an open-ended microstrip line fabricated on the silicon substrate – described in [5, 6]. However, the use of such a device in the oscillator tuning circuit gives not enough good performance.

3. PIN photodiodes

3.1. PIN photodiode performance

The commonly used microwave semiconductor devices with possibility of the optical control are photodiodes – especially PIN type ones, and phototransistors.
A PIN photodiode is a semiconductor junction device consisting of *p*-type and *n*-type semiconductor regions separated by an intrinsic layer (*i*) (Fig. 4a). In normal mode of operation a reverse-bias voltage applied across the device is large enough to make the intrinsic region fully depleted of carriers. An electric field resulting from reverse bias is equal to the saturation field. It means that during most of the transit the carriers travel at the saturation velocity.
When the light illuminates the semiconductor and the photons have its energy equal to or greater than the semiconductor band gap the generation of free electron-hole pairs takes place. The high electric field present in the depletion region separates the photocarriers. As the carriers traverse the depletion region, the current flow is induced. It means that in the equivalent circuit of PIN photodiode the optical illumination effect can be model by current source [7].
The performance of a photodiode is often characterised by the responsivity R. The photodiode responsivity is a ratio of generated photocurrent I_{ph} to the optical illumination power P_{opt}:

$$R = \frac{I_{ph}}{P_{opt}} \tag{3}$$

and is expressed in A/W.

The responsivity can be expressed also with the use of internal quantum efficiency η:

$$R = \frac{\eta q}{hf}. \tag{4}$$

The internal quantum efficiency is the number of electron-hole pairs generated per incident photon of energy equal to hf. The typical values of commercially available PIN photodiodes are equal to about 0.8 A/W.

The optical illumination causes the change in the junction parameters and thereby changes the microwave scattering parameters S_{11} of the photodiode. Expect the photodetection, this effect can be used to perform an optical and a microwave signals mixing. Also the optical control of a microwave circuit containing the PIN photodiode can be done in this way. The examples of S_{11} measurements in the case of optical illumination are shown in Fig. 5.

The impedance changes of typical PIN photodiode under the optical illumination are relatively small and sometimes not sufficient to use in some microwave circuits applications [8, 9]. Therefore there are other ways to improve the possibility of the microwave circuit control by means of the optical illumination and with the use of the photodiodes. It is worth to point an approach using the photodiode set to work with a varactor diode [8]. This approach is used in the microwave phase shifters [8, 10].

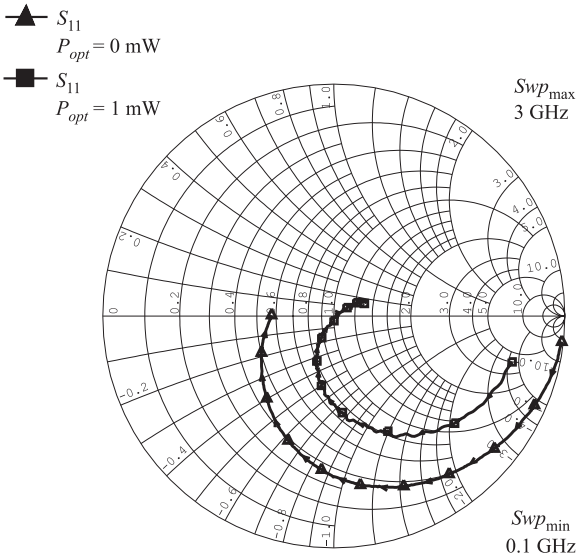


Fig. 5. Example of microwave reflectance of PIN photodiode under optical illumination.

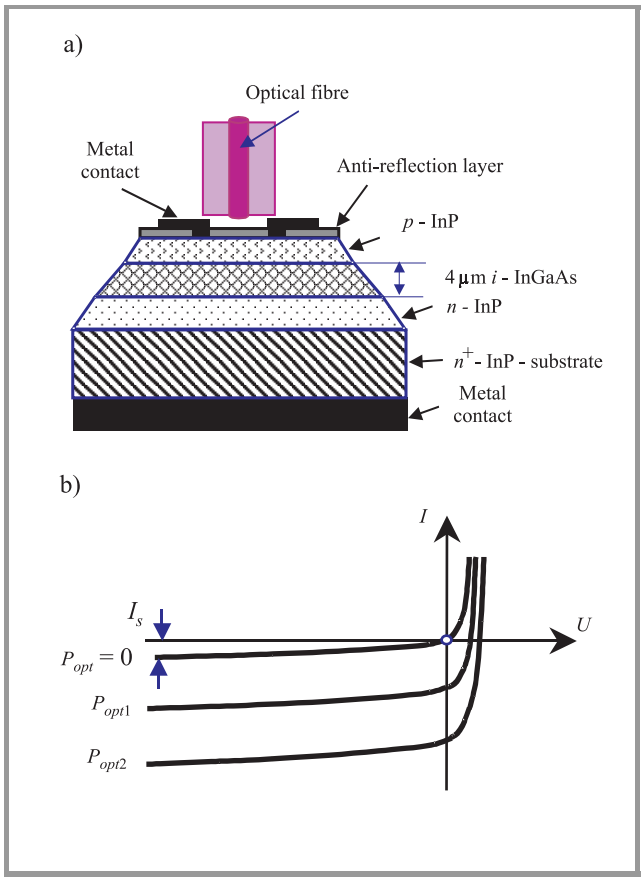


Fig. 4. PIN photodiode: (a) schematic cross-section of PIN photodiode; (b) diode characteristics $I(U)$ for different optical power.

3.2. Microwave parameters of PIN photodiode

A microwave behaviour of the PIN photodiode can be described by means of, either the scattering parameters, or the equivalent circuit/model. To determine the PIN photodiode equivalent circuit from the measurements, the special

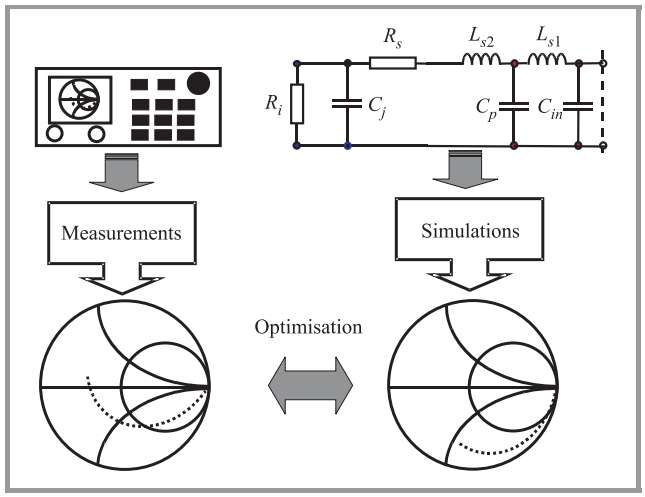


Fig. 6. Illustration of equivalent circuit extraction method from the measurements.

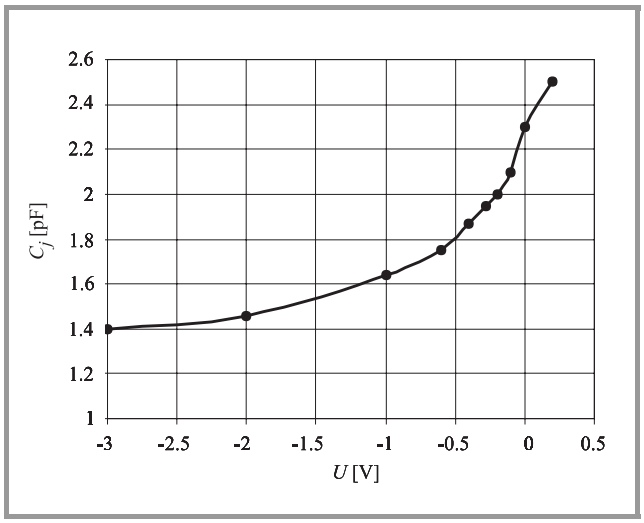


Fig. 7. Extracted PIN junction capacitance versus bias voltage characteristic.

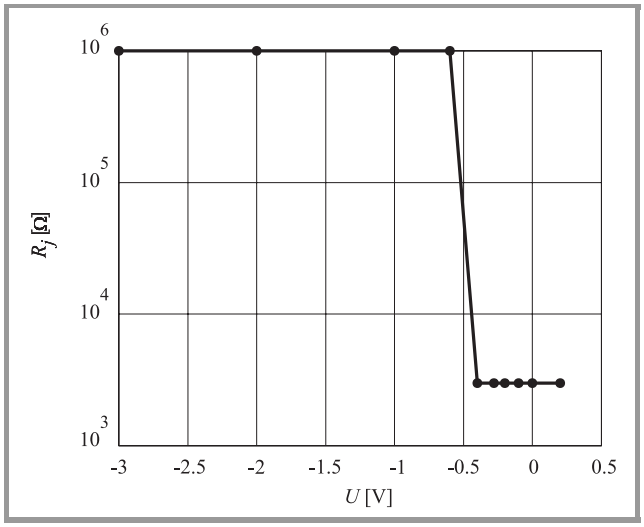


Fig. 8. Extracted PIN junction resistance versus bias voltage characteristic.

method of extraction was used. This method contains three steps:

a) measurements of the photodiode reflection coefficient Γ_{meas} within a defined frequency range, and at different bias voltages;

b) assumption of the equivalent circuit scheme, i.e. defining the R, L, C elements being contained;

c) the use of the special microwave simulating software (in our case *Microwave Office*) to calculate Γ_{sim}.

After these steps the final part of the extraction can be made. It is the optimisation of the variable values of the equivalent circuit elements to achieve fitting the simulation results to the measurements (Fig. 6). When matching is impossible, then the equivalent circuit is modified. The good matching, obtained within the specified frequency range, and at specified bias voltage, gives the values of the model parameters for these frequencies and at this voltage.

An example of PIN photodiode model extracted from the S_{11} measurements is shown below. The values of the parasitic elements are: $C_p = 0.03$ pF, $C_{in} = 0.4$ pF, $L_{s1} = 3$ nH, $L_{s2} = 4.1$ nH. The value of junction series resistance was found to be approximately constant and equal to $R_s = 29\ \Omega$. The extraction procedure repeated for different bias voltages gives the photodiode voltage characteristics (see Figs. 7 and 8).

4. Photovaractor

A type of semiconductor diodes very commonly used in the microwave circuits is a varactor diode. However, traditional varactor diode is the device that has the junction capacitance with possibility of steering by means of the bias voltage applied across the device. The semiconductor structure of the varactor is specially designed to obtain relatively high range of the capacitance changes while biasing at different voltages. A new approach to the problem of changing the capacitance in the microwave device or circuit is the use of the optical illumination. The photovaractor is the new special kind of the semiconductor PIN diode that has possibility of steering the junction capacitance with the use of the optical power [11]. The photovaractor structure is specially designed to obtain the capacitance changes as high as possible while illuminating with the optical power. A typical photovaractor diode is a heterostructure type device. An example of the photovaractor cross-section is shown in Fig. 9.

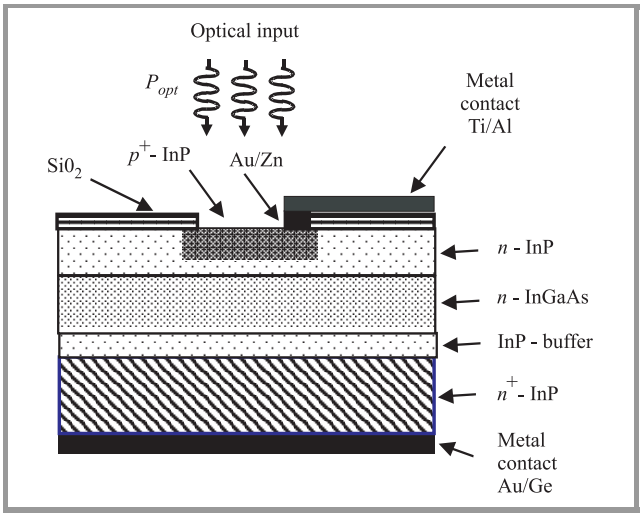


Fig. 9. Example of n-InP/n-InGaAs/n^+-InP photovaractor cross-section [7].

The maximal value of the photovaractor capacitance change under the optical illumination occurs at the bias around 0 V. For example, at the bias voltage equal to +0.1 V the capacitance changes from $C_{\min} = 2.75$ pF up to $C_{\max} = 69$ pF, under the optical power variation from 0 up

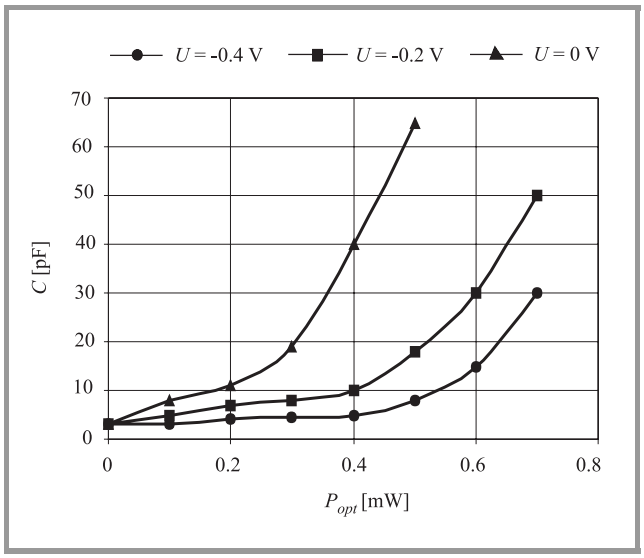


Fig. 10. Measured photovaractor capacitance versus optical power characteristics at different bias voltages.

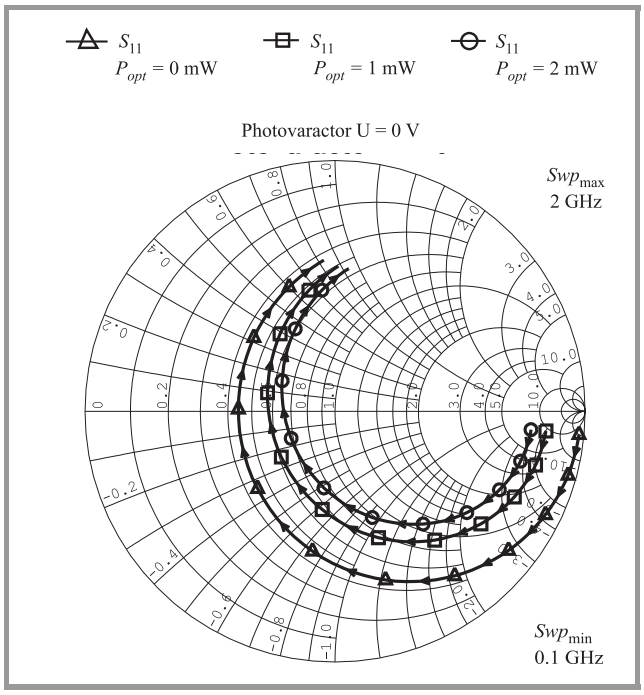


Fig. 11. Example of microwave reflectance S_{11} of PIN photovaractor under optical illumination.

to 0.4 mW. Then the ratio $C_{\max}/C_{\min}$ is equal to 25. An example of the photovaractor measured capacitance versus the optical power characteristics at different bias voltages is shown in Fig. 10.

However, the highest value of capacitance changes can be noticed for the capacitance measured at the several hundreds of MHz. The capacitance measured or extracted at the higher frequencies, for example few GHz, has smaller values because of fact that at high frequencies the generated photocarriers have no time to follow the change of a microwave signal. Nevertheless, the photovaractor capacitance at the microwave frequencies can be sufficient to provide the optical control of various microwave circuits, especially the oscillators (Fig. 11) [12, 13].

5. Optically variable capacitor

5.1. Photodiode-varactor couple

An optically variable capacitor (OVC) is new type of microwave-optical device. This device is the connection of the photodiode and the varactor diode. The photodiode is placed with a proper bias circuit, which allows to change the bias voltage across the varactor by means of the optical power illuminating the photodiode. The change of the bias voltage causes the change of the varactor capacitance, thereby allowing optical control of the circuit containing the varactor (Fig. 12). The problem of the connection can be solved in different ways, and two of the solutions are specially interesting.

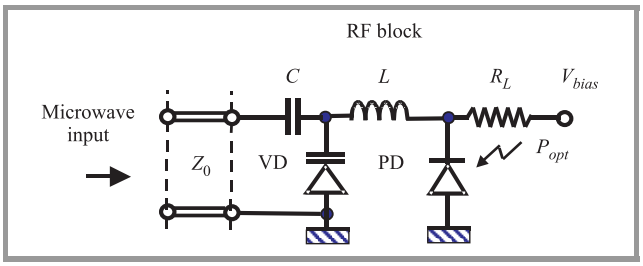


Fig. 12. Photodiode-varactor couple as a microwave optically variable capacitor.

The maximum of the optical power illuminating the photodiode corresponds to minimum of bias voltage across the varactor.

Example of microwave performance of the OVC is shown in Fig. 13. There is presented the OVC reflection coefficient S_{11} for different values of the optical power from 0 to 3 mW at the frequency equal to 1 GHz.

This type of the OVC circuit has the following features:

- RF block inductance keeps photodiode out of microwave signal;
- reverse biased varactor dissipates very little power;
- optical power required for control depends on the resistor used in the bias circuit, and can be relatively small;
- optical and microwave functions are performed in separate devices and they can be independently optimised;
- varactor diode used in this OVC can be assort with respect to highest capacitance changes ratio.

This type of circuit was developed to investigate the possibility of obtain an optically controlled microwave device for special microwave phase shifters and modulators [8, 10].

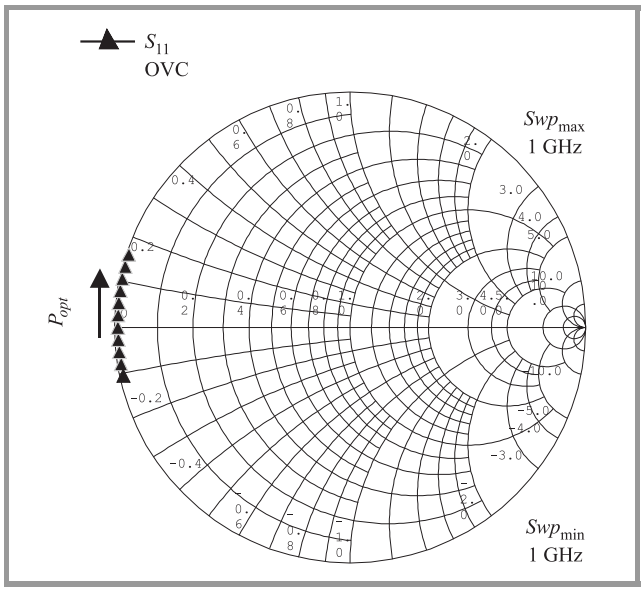


Fig. 13. OVC reflection coefficient S_{11} for different values of the optical power from 0 to 3 mW at the frequency equal to 1 GHz.

5.2. Optically variable capacitor with self-bias

Another very interesting solution is the device shown in Fig. 14. This device is the connection of photovoltaic cell with the varactor diode. The optical illumination of photovoltaic cell array causes photocurrent generation and then this current converted into the voltage is used to bias the varactor. In this manner the optical control of the varactor capacitance is provided. It is important to notice that this OVC is a bias-free device [14].

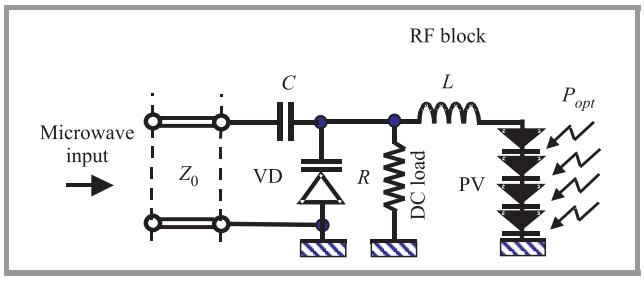


Fig. 14. Schematic idea of optically variable capacitor with self-bias using photovoltaic cell.

This OVC structure offers several advantages:

- RF block inductance keeps PV array out of microwave signal;
- reverse biased varactor dissipates very little power;
- optical power required for control is small;
- optical and microwave functions are performed in separate devices and they can be independently optimised;
- varactor diode is designed to produce desired capacitance swing with lowest possible RF insertion loss;

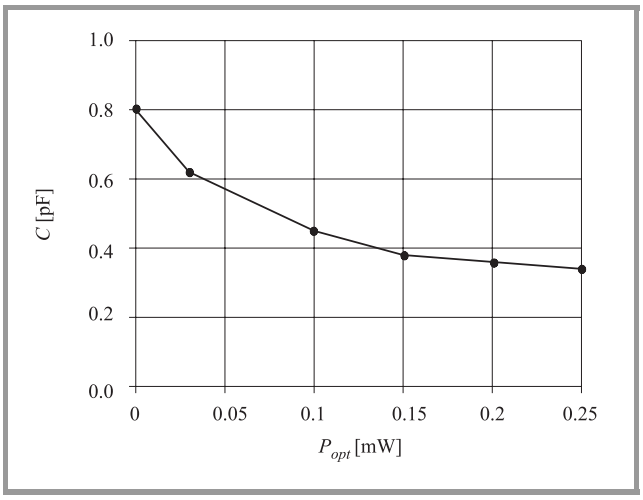


Fig. 15. Example of OVC performance: varactor capacitance versus optical power.

- photovoltaic call array is designed to generate desired output voltage range using the smallest optical power.

The comparison of two solutions described in this Section allows to conclude that the solution with the PIN photodiode provides faster change of the varactor capacitance, because the working bandwidth of the photovoltaic cells is usually small (Fig. 15).

6. Phototransistors

6.1. General types of phototransistors

The transistors are the types of semiconductor devices sensitive on the optical illumination. There are three types of microwave transistors: MESFET, HEMT and HBT [15 – 17]. Their main application is the photodetection. As the photodetectors transistors offer several advantages:

- detected signal is amplified by the device transconductance;
- in the case of HEMT the improvement of output signal to noise ratio;
- possibility of use of semiconductor materials with high electron mobility – like AlGaAs, InGaAs, InGaAsP to form the heterostructures in HEMT transistors; the energy gap in these materials can be chosen to obtain maximal sensitivity for given light wavelength;
- possibility of integration with other optoelectronic devices (even planar optics).

Generally, the optical illumination of the transistor causes the change of its parameters. This phenomenon, beside photodetection, can be used for mixing of the microwave and the optical signals. Because the optical illumination

changes the microwave scattering parameters (S matrix) of the transistor then it is possible to optically control the microwave oscillator containing this device. Typical structures of microwave transistors MESFET, and HBT were shown in Figs. 16 and 17 [15, 16].

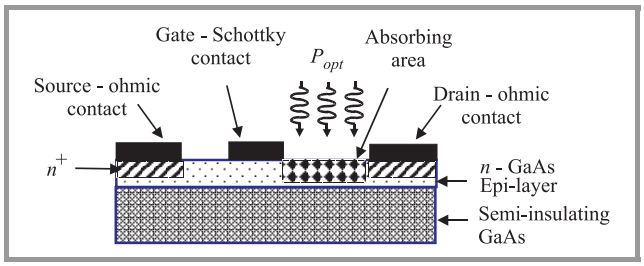


Fig. 16. Schematic cross-section of typical MESFET GaAs phototransistor structure.

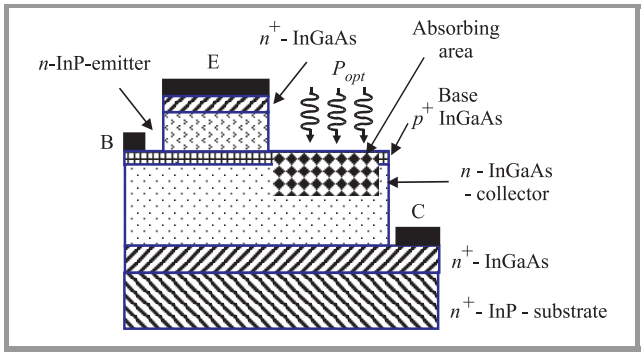


Fig. 17. Typical structure of HBT phototransistor.

From the technological point of view, the phototransistor structures are similar to the classical transistor structures. Nevertheless, there are two general requirements to fulfil. First, the light reflection at the transistor surface has to be as small as possible with the good transmission at the same time. Second, the light absorption has to be very high in the semiconductor layer chosen for this purpose. It is the channel area in the FET type transistors and the base-collector junction in case of bipolar junction transistors (BJT).

Generally the optical illumination of a semiconductor material causes the free electron-hole generation process. In the case of FET transistors this process increases the concentration of the minority carriers, for example, the holes in a n-type channel. Mathematically, Δp is expressed as:

$$\Delta p = \frac{\tau}{d} \cdot \frac{P_{opt}\lambda}{hc} \cdot \left[1 - \exp(-\alpha d)\right], \tag{5}$$

where P_{opt} is an incident optical power per unit area, d is a thickness of the active layer, τ is a minority carrier lifetime and c is the speed of light in vacuum.

The consequence of the excess holes concentration is the light-induced voltage V_{lit} at the Schottky gate:

$$V_{lit} = \frac{kT}{q} \ln\left[\frac{p + \Delta p}{p}\right], \tag{6}$$

where T is a temperature in Kelvins and q is an electron charge.

Hence, the illumination of the gate region has the same effect as forward biasing the gate.

The optical performance of phototransistors is also characterised by the responsivity R, which is expressed as

$$R = \frac{I_p}{P_{opt}}, \tag{7}$$

where I_p is the drain current under the illumination with the power P_{opt}.

The example of drain-source current I_{DS} versus drain-source voltage U_{DS} characteristics at different values of gate-source voltage U_{GS} and the illumination power is shown in Fig. 18.

In the case of bipolar transistors BJT the simplified modelling is used very often. These models are based on the assumption that the structure of a transistor, for example HBT, is configured to obtain maximal absorption of the light in the area of the base-collector junction.

Because in the active mode of operation the base-collector junction is reverse biased and the absorption of light takes place mainly in the junction depletion layer, the photocurrent flow is generated. Thus, the optical illumination of the transistor can be modelled with the use of additional

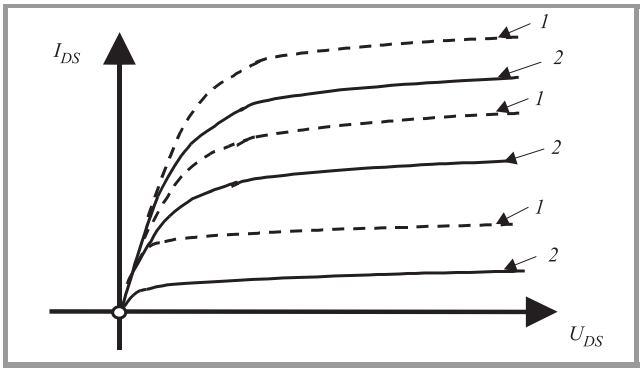


Fig. 18. Example of drain-source current I_{DS} versus drain-source voltage U_{DS} characteristics for HEMT with (*1*), and without (*2*) optical illumination.

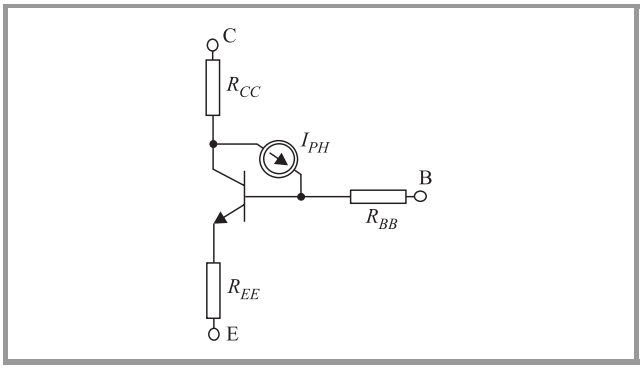


Fig. 19. Simplified SPICE-based HBT phototransistor model.

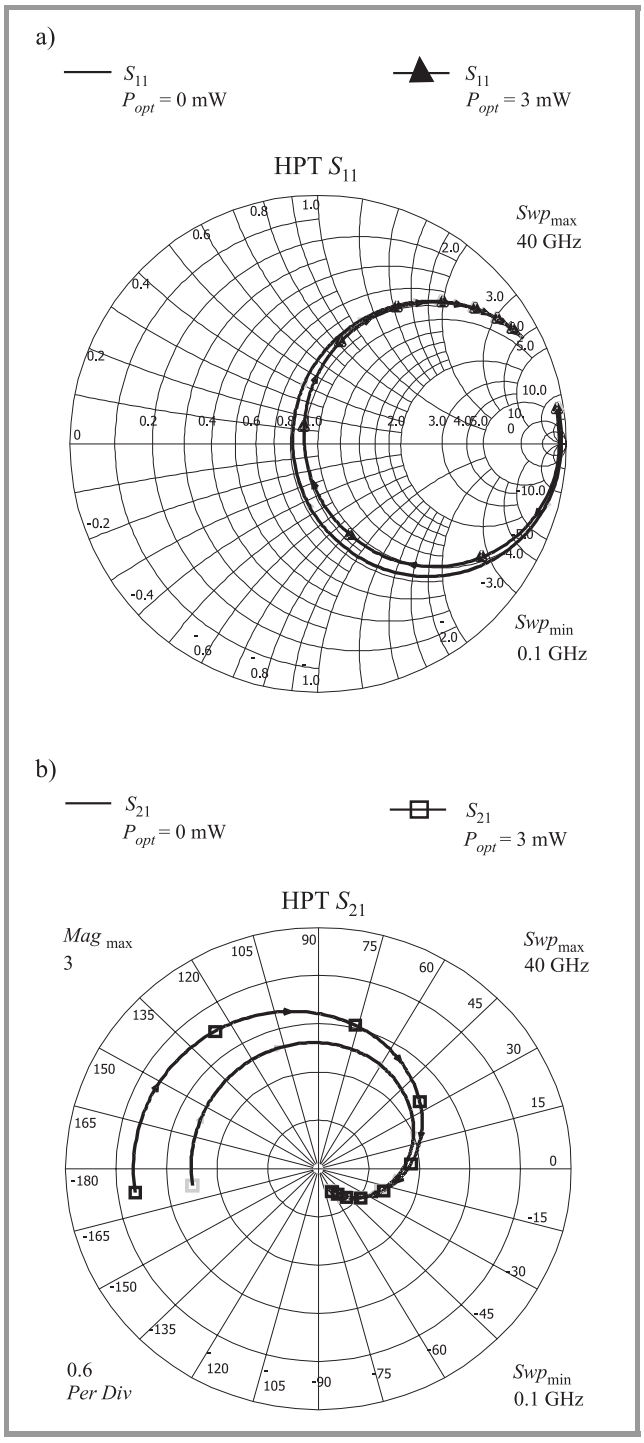


Fig. 20. Scattering parameters S_{11} (a) and S_{21} (b) characteristics respectively for 0 mW, and 3.0 mW of optical illumination.

current source connected between the base and the collector terminals. The value of photocurrent depends on the detection quantum efficiency of the base-collector junction. The quantum efficiency can be improved by increasing of the junction (depletion layer) thickness. On the other hand the light penetration depth in the semiconductor material should be approximately equal to the thickness of the base and collector layers.

6.2. Simulation of HBT phototransistor performance

The phototransistor parameters without illumination can be described with the use of standard modelling – for example the SPICE type parameters. There are other models available in the special simulating microwave software, for example Gummel-Poon. Then the additional current source can be used to consider the illumination influence (Fig. 19).

Such an improved model allowed to simulate the microwave scattering parameters of the phototransistor configured in common emitter. The knowledge of the HPT *S*-parameters microwave performance allows to design the microwave circuits – such an oscillator, or a mixer – containing the HPT device.

The examples of HBT phototransistor behaviour under the optical illumination are shown in Fig. 20. There are shown the scattering parameters S_{11}, S_{21} for the collector-emitter bias $U_{CE} = 2$ V and the base-emitter bias $U_{BE} = 0.77$ V, and the frequency band from 0.1 to 40 GHz.

All the *S*-parameters of the phototransistor depend on the optical illumination. Hence, there is possibility of use of the phototransistor as an optically controlled active device in the microwave oscillators, mixers and amplifiers [1, 15 – 17].

7. Conclusions

In many applications there is a need to control the microwave circuits performance and parameters by means of the optical illumination. For purposes of such a problem the photodevices which detect the optical signals and work in the optical receivers can be used. There are the PIN photodiodes and the phototransistors in this group of the devices. On the other hand the new types of the devices were developed, such as photoconductive switches, photovaractors or optically variable capacitors. These devices can be used in many different microwave circuits such oscillators, phase shifters, attenuators or ultra-short electrical pulse sources. There are many devices, which have relatively good performance. However, there are some problems, which are not overcome yet. The main problem of all microwave photonic devices is to achieve very large microwave bandwidth, with high optical quantum efficiency simultaneously. The next generations of the devices show the growing development in this field.

Acknowledgements

This work was supported by the State Committee for Scientific Research under contract no. 8T11D04214.

References

[1] B. A. Galwas, "Photonic technology for microwave engineering", in *12th Int. Microw. Conf. (MIKON'98)*, Cracow, Poland, May 1998.

[2] I. A. McIntyre, D. M. Giorgi, D. E. Hargis, and O. S. F. Zucker, "Optical requirements for light activated switches", *Proc. SPIE*, vol. 1378, pp. 162–172, 1990.

[3] S. E. Saddow and Ch. H. Lee, "Optical control of microwave-integrated circuits using high-speed GaAs and Si photoconductive switches", *IEEE Trans. MTT.*, vol. 43, no. 9, pp. 2414–2420, 1995.

[4] X. Xue, H. Wei, and I. Shih, "Optically controlled switching in microwave bridges", in *Proc. 30th Eur. Microw. Conf. (EuMC'2000)*, Paris, France, Oct. 2000.

[5] B. Cabon *et al.*, "Optical control of microwave passive devices integrated on silicon", in *Proc. 10th Conf. MICROCOLL'99*, Budapest, Hungary, March 1999.

[6] P. Ahouassa *et al.*, "Microwave photoinduced varactor on silicon substrate", in *Proc. 30th Eur. Microw. Conf. (EuMC'2000)*, Paris, France, Oct. 2000.

[7] A. L. Chizh and S. A. Malyshev, "Modeling and characterization of microwave p-i-n photodiode", in *Proc. Third Int. Conf. Adv. Semicond. Dev. Microsyst. (ASDAM'2000)*, Smolenice Castle, Slovakia, Oct. 2000, pp. 239–242.

[8] Z. Szczepaniak and B. Galwas, "Optically-controlled microwave phase shifter", in *Proc. Conf. EDMO'99*, London, U.K., Nov. 1999.

[9] Z. Varga, G. Jaro, and B. Erdos, "Optical control of microwave filters using photodiodes", in *Proc. 10th Conf. MICROCOLL'99*, Budapest, Hungary, March 1999.

[10] Z. Szczepaniak and B. Galwas, "Optically-controlled microwave PSK modulator", in *Int. Microw. Conf. (MIKON'2000)*, Wroclaw, Poland, May 2000.

[11] S. Malyshev and A. Chizh, "Modeling and characterization of photovaractor for microwave optoelectronics", in *Proc. 31th Eur. Microw. Conf. (EuMC'2001)*, London, U.K., Sept. 2001, vol. 1, pp. 221–224.

[12] Z. R. Szczepaniak, B. A. Galwas, and S. A. Malyshev, "Optically-switched microwave filter with the use of photovaractors in self-bias mode", in *Proc. 31th Eur. Microw. Conf. (EuMC'2001)*, London, U.K., Sept. 2001.

[13] Z. R. Szczepaniak and B. A. Galwas, "The use of photovaractor for optical control of microwave oscillator frequency", in *Proc. 11th Conf. COMITE'2001*, Pardubice, Czech Republic, Sept. 2001.

[14] A. S Nagra *et al.*, "Indirect optical control of microwave circuits using monolithic optically variable capacitors", *IEEE Trans. MTT*, vol. 47, no. 7, pp. 1365–1372, 1999.

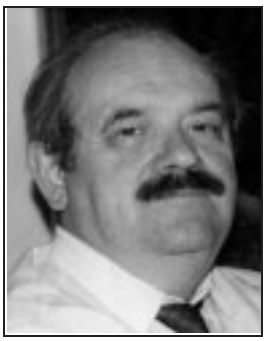
[15] T. Shimizu, M. Nakatsugawa, and H. Othsuka, "Performance of GaAs MESFET photodetectors with wide drain-to-gate distances in subcarrier optical transmission", *IEICE Trans. Electron.*, vol. E80-C, no. 1, pp. 160–167, 1997.

[16] Y. Betser *et al.*, "A single-stage three-terminal heterojunction bipolar transistor optoelectronic mixer", *J. Lightw. Technol.*, vol. 16, no. 4, pp. 605–609, 1998.

[17] P. N. Freeman *et al.*, "Optically controlled microwave oscillators fabricated using GaAs/AlGaAs HBTs with transparent ITO emitter contacts", *Proc. SPIE*, vol. 2155, pp. 24–28, 1994.

Zenon R. Szczepaniak received the M.Sc. degree in optoelectronics from Warsaw University of Technology (WUT), Poland in 1998. Since 1998 he is a Ph.D. student in Microwave Devices Division at the Warsaw University of Technology. His research interests include optically-controlled microwave circuits, modelling of optical-microwave devices and systems for telecommunication.

e-mail: zenon.szczepaniak@elka.pw.edu.pl
Institute of Microelectronics and Optoelectronics
Warsaw University of Technology
Koszykowa st 75
00-662 Warsaw, Poland



Bogdan A. Galwas was born in Poland, on October 31, 1938. In 1962, he joined the Faculty of Electronics Warsaw University as Lecturer. He received the M.Sc. degree in 1962, the Ph.D. degree in 1969, and the D.Sc. degree in 1976, all in electronic engineering from Warsaw University of Technology, Poland. In 1986 he was promoted to Full Professor. His current research interests are microwave electronics and photonics. He is the author of more than 120 scientific papers and 2 books in these areas. His main field of academic interest is connected with technology of education, continuing engineering education and open distance learning. He is a Chairman of the International Management Committee of the International Travelling Summer Schools '91, Member of IACEE '97 and Member of SEFI '97.

e-mail: galwas@imio.pw.edu.pl
Institute of Microelectronics and Optoelectronics
Warsaw University of Technology
Koszykowa st 75
00-662 Warsaw, Poland

ICTON 2002

4th International Conference on Transparent Optical Networks

Co-located with: European Symposium on Photonic Crystals ESPC 2002

CALL FOR PAPERS

The National Institute of Telecommunications in Warsaw, Department of Transmission and Fibre Technology, and IEEE/LEOS Poland Chapter are pleased to announce the **4th International Conference on Transparent Optical Networks ICTON 2002**, which will be held at the National Institute of Telecommunications, 1 Szachowa Street, Warsaw, Poland, **April 21-25, 2002**. ICTON 2002 will be technically co-sponsored by the IEEE Lasers and Electro-Optics Society. European Projects COST P2: "Applications of Non Linear Optical Phenomena" and COST 266: "Advanced Infrastructure for Photonic Networks" are associated with the organisation of the Conference. The **European Symposium on Photonic Crystals ESPC 2002** will be organised as a part of ICTON 2002 in association with COST 268: "Wavelength Scale Photonic Components for Telecommunications" Project. The ESPC is intended to cover recent technology and theory & modelling achievements within the field with emphasis on optical signal processing for telecommunication applications.
An **exhibition** of photonic telecommunication industry is planned with the Conference.

Warsaw which is the Poland's capital city is located in the centre of the country with convenient road, rail and air connections to the European network. The central position in Europe as well as an international airport makes the city easily accessible.

The scope of the Conference is concentrated on the applications of all-optical technology in transparent networks, systems, and components, and includes:

- Terabit Transport Networks
- Ultra-dense WDM transmission
- Optical time domain multiplexing
- Wavelength conversion
- Soliton transmission
- Dispersion management and compensation
- Polarisation Mode Dispersion
- Optical amplifiers
- New fibre types
- New transmission windows
- New light sources
- Fibre gratings
- Arrayed waveguide gratings
- **Photonic Band-Gap structures (ESPC)**
- VCSELs
- All-optical signal regeneration
- All-optical signal processing
- Optical switching
- Optical memory and data storage
- Non-linear waveguide optics
- Analogue transmission systems
- Broadband access networks
- Optical transmission of microwaves
- Modelling of optical systems and components
- Standardisation of optical networks
- IP over WDM
- Network migration

Authors are cordially invited to send their **contributions** (4 pages in electronic form, Word 6.0 or later version accompanied with PDF version) for Regular Sessions and Poster Session to the Organising Committee by **February 28, 2002**. The authors will be notified on the acceptance by March 15, 2002. **Post-deadline contributions** with very recent results are solicited until **April 2, 2002**.

ICTON Organising Committee:
Marian Marciniak, Doctor of Sciences – Chairman
National Institute of Telecommunications
Department of Transmission and Fibre Technology
Szachowa 1, 04-894 Warsaw, Poland
Tel. (+48 22) 812 00 72 Fax (+48 22) 5128 347
E-mail: ICTON@ITL.WAW.PL , http://www.itl.waw.pl/icton/

ICTON Scientific Committee:
Prof. Andrzej P. Wierzbicki – Chairman
National Institute of Telecommunications
Szachowa 1, 04-894 Warsaw, Poland

National Institute
of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

Editorial Office
tel. +48(22) 872 43 88
tel./fax:+48(22) 512 84 00
e-mail: redakcja@itl.waw.pl
http://www.itl.waw.pl/jtit