Coded Orthogonal Frequency Division Multiplex

BERNARD LE FLOCH, MICHEL ALARD, MEMBER, IEEE, AND CLAUDE BERROU, MEMBER, IEEE

Technological evolution and the ever-increasing demand for higher-quality services give broadcasters a strong incentive to completely digitize their broadcasting networks. This digitization, which is already well advanced in many program production areas and transmission links, now has to be extended to complete the last link in the broadcast chain; i.e., from broadcast transmitter to consumer receivers.

It is therefore necessary to develop wholly new techniques for the broadcasting of digitally coded TV programmes. Thus an efficient baseband digital coding must be combined with a robust digital modulation and channel coding scheme that can meet the requirements of every mode of broadcast reception.

This article presents the research work related to the coded orthogonal frequency division multiplex (COFDM) technology, which has now been completed in the field of digital radio (DAB) [1], and which is under progress in the field of digital terrestrial TV.

I. Introduction

Terrestrial broadcasting is faced with the problems of echoes due to the phenomena of multiple propagation, and saturation of the spectral resource. The intrinsic quality of the signal, and the frequency planning, are two aspects which are often considered as being independent. However the interference caused by long distance transmitters broadcasting the same program on the same frequency as the local transmitter can be considered as artificial echoes, and consequently the two aspects underlined above are in fact related to the same issue. The potentiality of a new broadcasting system will therefore depend on the strategy adopted regarding the echoes, whether they are natural or artificial. If the way of dealing with echoes is innovative and systematically takes advantage of multipath propagation instead of enduring it, the new perspectives opened up by the digital system are really revolutionary.

The following article goes over the principles of a multicarrier modulation scheme developed with this goal in mind. This scheme was initially designed and adopted for the European Digital Audio Broadcasting project, but is equally suitable for television broadcasting. This article also highlights its advantages and gives prominence to the beneficial consequences (through the use of a single frequency network (SFN)) to the economy and engineering of terrestrial TV broadcasting networks.

Section II gives an overview of the principles of coded multicarrier transmission systems. Section III, which covers a broader search for an optimum OFDM signal shaping, gives a prospective analysis of the possible options for defining the orthogonal basis of elementary symbols forming the transmitted signal: The original signal configuration using a rectangular time-window shaping with guard interval, as well as a new time-frequency shaping concept called isotropic orthogonal transform algorithm (IOTA). Section IV deals with the choice of the channel coding technique. Section V provides performance results as a function of the tradeoff between power and spectrum efficiencies under different transmission channel conditions, and Section VI gives a comparison of performances between COFDM and linear equalization of single carrier transmission. Section VII stresses the advantages of COFDM in terms of its suitability with respect to broadcast engineering and spectrum management.

II. General Principles of COFDM

The COFDM technique is particularly suited to provide reliable reception of signals affected by strong distortions, as it can be the case for terrestrial broadcasting [2]. Multipath propagation is known to limit the performance of high bit-rate modulation schemes. The principle of COFDM relies in splitting the information to be transmitted over a large number of carriers, in such a way that the signaling rate on each of them becomes significantly lower than the assumed channel coherence bandwidth. In other words, the signal is conditioned to ensure that modulated symbols will be much longer than the echo delay spread. Provided that a guard interval is inserted between successive symbols, multipath propagation does not generate anymore intersymbol interferences.

However, in presence of strong echoes, some carriers will suffer deep fades, due to destructive combination of the various reflections, while others will be enhanced by constructive addition. The signal-to-noise ratio at the receiver input increases as soon as the signal power is augmented by echoes separated by at least a delay equal to the inverse of the signal bandwidth. To benefit from this

Manuscript received July 1, 1994, revised March 2, 1995.

B. Le Floch is with the Centre Commun d'Etudes de Télédiffusion et Télécommunications (CCETT), 35512 Cesson Sévigné Cedex, France.

M. Alard is with WAVECOM, 92130 Issy-les-Moulineaux, France.

C. Berrou is with the Ecole Nationale Supérieure des Télécommunications de Bretagne (ENST Br), 29285 Brest Cedex, France.

IEEE Log Number 94011374.

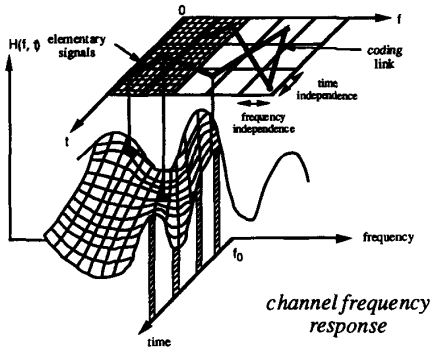


Fig. 1. Principles of COFDM.

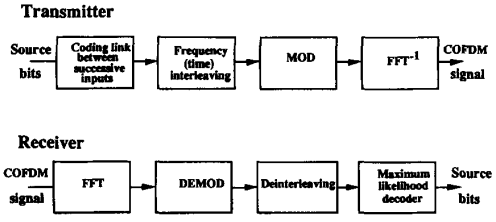


Fig. 2. Block diagram of a COFDM modem.

power increase, despite the fact that parts of the spectrum are deeply faded, it is necessary to incorporate a powerful channel coding scheme in the system design [3].

The role of coding, in conjunction with frequency and time interleaving, is to provide a link between bits transmitted on separated carriers of the signal spectrum (see Fig. 1), in such a way that information conveyed by faded carriers can be reconstructed in the receiver, thanks to the coding link which relates it to information conveyed by well-received carriers. Coding and interleaving applied to OFDM can be seen as a tool to average local fadings over the whole signal bandwidth and over the time interleaving depth. Frequency selectivity, currently known to be a disadvantage, is then turned into an advantage that can be called frequency diversity. This is the key feature which allows for SFN operation.

The block diagram of the functions included in a COFDM modem are presented in Fig. 2. It is important to notice that the tool used to build the signal at the transmitter and to analyze it at the receiver is the discrete Fourier transform [4], which allows for low cost implementation due to the existence of Fast algorithms.

In the classical implementation of COFDM, the elementary symbols constituting the signal are built of time-limited slices of sine and cosine waveforms. The transmitted signal can therefore be written in the following analytic form

$$s(t) = \sum_{n=-\infty}^{+\infty} \sum_{m=0}^{M-1} a_{m,n} x_{m,n}(t) \tag{1}$$

where $a_{m,n}$ represents one of the elements of the modulation alphabet conveyed by the carrier of index m during the symbol time of index n, and $x_{m,n}(t)$ represents the generic function of the signal basis

$$x_{m,n}(t) = e^{2i\pi mt/\tau_0}.\mathrm{Rect}(t - nT_0). \tag{2}$$

The signaling duration T_0 is chosen to be longer than the time interval τ_0 during which the carriers are orthogonal (the carrier distance is equal to the inverse of τ_0). This duration T_0 is the sum of the guard interval Δ and the "useful symbol time" τ_0.

The Rect function is defined by

$$\mathrm{Rect}(t) = \begin{cases} 1/\sqrt{\tau_0} & \text{if } -T_0/2 \le t < T_0/2. \\ 0 & \text{elsewhere} \end{cases} \tag{3}$$

Under the realistic assumption that the cumulated contributions of the secondary lobes of the carriers spectra, arising from the time-limited symbol shaping, is negligible outside the interval $[0, M/\tau_0]$, which is the case because a limited number of carriers on the spectrum edges are not transmitted (in order to minimize adjacent channel interference), it is possible to construct the signal from its samples $s(k\tau_0/M)$, where

$$s(k\tau_0/M) = \sum_{m=0}^{M-1} a_m e^{2i\pi mk/M}. \tag{4}$$

In this equation, which holds for any COFDM symbol, the time index n has been omitted without any loss of generality.

Therefore, $s(k\tau_0/M)$ is obtained by the discrete inverse Fourier transform of $\{a_m\}_{m=0}^{M-1}$.

Under the assumption that the echoes remain shorter than the guard interval, the received signal observed over the "useful symbol time" τ_0 is affected by intrasymbol interference only. This means that the modulation symbols a_m are transformed by the channel into $H_m a_m$, where H_m is the channel frequency response at frequency index m. These terms $H_m a_m$ are recovered in the receiver using a discrete direct Fourier transform; if $s'(t)$ is the received signal

$$H_m a_m = \sum_{k=0}^{M-1} s'(k\tau_0/M) e^{-2i\pi mk/M}. \tag{5}$$

In order to achieve a coherent demodulation, the channel frequency response H_m must be recovered for each value of m. This is generally achieved by inserting in the transmitted signal a grid of time-frequency spread reference pilots [5], which comply with the sampling theorem with regard to the frequency and time dispersion of the channel. Interpolation applied on these references allows estimation of the channel. The term H_m includes a phase rotation and a modulus (ρ_m), representing the attenuation factor of the channel at a given frequency: Some carriers are enhanced by constructive interference, while others are faded by a destructive combination of the echoes. After derotating the samples at the Fourier transform output, the terms $\rho_m a_m$ are processed by the channel decoder (desinterleaving and maximum likelihood decoding), in which the attenuation factor ρ_m is also taken into account as a measure of the

reliability of the information conveyed by the carrier of index m.

III. Discussion on a Broader Search for an Optimum OFDM Orthogonal Basis

This section constitutes a prospective analysis of the optimality of the choice of the signal basis in order to improve the system performance for the most difficult channel conditions. As only the construction of this basis is discussed hereunder and hence the discussion on the coding is not included, the letter C of COFDM is generally omitted.

A. Basic Principles

1) OFDM and Hilbertian Basis: We consider first the case of an OFDM signal $s(t)$ without any guard interval

$$s(t) = \sum_{m,n} a_{m,n} x_{m,n}(t) \tag{6}$$

where the coefficients $a_{m,n}$ take complex values representing the transmitted encoded data, and the basic functions $x_{m,n}(t)$ are obtained by translation in time and frequency of a prototype function $x(t)$

$$x_{m,n}(t) = e^{2i\pi m\nu_0 t} x(t - n\tau_0) \text{ with } \nu_0\tau_0 = 1 \tag{7}$$

and

$$x(t) = \begin{cases} \frac{1}{\sqrt{\tau_0}} & \text{if } |t| \le \tau_0/2 \\ 0 & \text{elsewhere} \end{cases} \tag{8}$$

We can easily give a graphical representation of this set of functions in the time-frequency plane by their first order moments

$$\int t\|x_{m,n}(t)\|^2\, dt = n\tau_0$$
$$\int f\|X_{m,n}(f)\|^2\, df = m\nu_0. \tag{9}$$

The coordinates of the basic functions form a two-dimensional (2D) lattice in the time-frequency plane. The density of this lattice is equal to unity, that is $\nu_0\tau_0 = 1$. This set of functions is orthonormal and form a Hilbertian basis of $\mathbf{L}^2(\mathbf{R})$. It must be noted that $\nu_0\tau_0 = 1$ is a necessary condition (while not sufficient) in order to obtain a Hilbertian basis [7]. The projection of a given signal on this basis can be thought as partitioning the signal in subsequences of duration τ_0, each of them being represented by the associated Fourier series. This is a first step toward localization in both time and frequency domain, compared with conventional Fourier analysis, which provides a perfect frequency localization, but on the other hand loses the time information.

2) Hilbertian Basis and Modulation Schemes: The interest in Hilbertian basis for digital modulation can be understood from the Shannon theorem, which can be written as

$$\text{number of bits per dimension} = (\log_2(1+S/N))/2. \tag{10}$$

It is a well known result that any signal of duration T and bandwidth W can be represented, at least asymptotically, by $2WT$ dimensions. Therefore, it is an essential property of a modulation system to make the best use of all these available space dimensions.

Hilbertian basis provides a powerful tool to design modulation systems which are optimum from this point of view. The transmitted signal can be written as

$$s(t) = \sum a_k x_k(t) \tag{11}$$

where $\{x_k\}$ is a set of functions being a Hilbertian basis of $\mathbf{L}^2(\mathbf{R})$, and a_k being real or complex values representing the transmitted data. As $\{x_k\}$ is a basis of $\mathbf{L}^2(\mathbf{R})$, no space dimension is lost. Furthermore, if $\{x_k\}$ is a Hilbertian basis, the distance properties of the coding scheme are not changed by the modulation process. This is an essential property in a channel corrupted by gaussian noise, where the performances are directly related to the Euclidian distance between transmitted signals. Last, but not least, a Hilbertian basis is also desirable because it greatly simplifies the receiver design, for we can simply recover the transmitted data by taking benefit of the orthogonality of the basic functions

$$a_k = \langle s \mid x_k \rangle = \int_R s(t) x_k^*(t)\, dt. \tag{12}$$

A last criterion for designing such an Hilbertian basis is related to the transmission channel itself. In a dispersive channel, the transmitted signal is spread in time and in frequency. In practice, we can also include in the channel model other distortions of the signal due to the receiver implementation, such as local oscillator offset and phase noise, synchronization errors or jitter. We expect the transmitted signals to be distorted as little as possible in such a channel. The optimum solution to meet this requirement is to use basic signals which are localized in time and in frequency with the same time-frequency scale as the channel itself. More precisely, if $\Delta\tau$ and $\Delta\nu$ are respectively the channel delay spread and frequency spread, we can define a time frequency scale τ_0 and ν_0 so that

$$\frac{\tau_0}{\Delta\tau} = \frac{\nu_0}{\Delta\nu} \text{ and } \nu_0\tau_0 = 1. \tag{13}$$

3) About the Existence of Other Hilbertian Bases: The OFDM scheme previously described seems so straightforward that we could reasonably expect that other bases might exist, based on a similar approach. Another example is obtained by inverting the time and frequency axes. In this case, the prototype function is the Fourier transform of the initial window, i.e., a $(\sin x)/x$ function. This is in fact the asymptotic case of a system using nonoverlapping frequency bands for elementary channels, which can be separated by rectangular filters. We could call this approach the "zero roll-off approach." Unfortunately, other extensions are not obvious. Even though other functions having the same orthogonality properties as the rectangular window and the $(\sin x)/x$ window have been described in the past [6], they are not satisfactorily

localized in the time and frequency domains, and, in addition, they are asymmetrical. Furthermore the Balian–Low–Coifman–Semmes theorem [7] shows that if $x(t)$ is a prototype function having the same orthogonality properties as the rectangular window, then $\Delta t \Delta f = \infty$, Δt and Δf being the second order moments of the prototype function, defined as

$$\begin{cases} \Delta t^2 = \int t^2 \|x(t)\|^2 \, dt \\ \Delta f^2 = \int f^2 \|X(f)\|^2 \, df. \end{cases} \tag{14}$$

Therefore, we cannot expect to find a prototype function so that both $x(t)$ and $X(f)$ decrease faster than $|t|^{-3/2}$ and $|f|^{-3/2}$ respectively. This does not preclude that "good" prototype functions might exist from the engineering point of view, but this leaves in our mind few chances to find them.

Another classical basis of $\mathbf{L}^2(\mathbf{R})$ having good properties regarding time and frequency localization was introduced by Gabor [8]. The prototype function is a normalized gaussian function

$$x(t) = \frac{2^{1/4}}{\sqrt{\tau_0}} e^{-\pi t^2/\tau_0^2}. \tag{15}$$

It can be demonstrated that the set of functions $x_{m,n}(t)$ forms a basis, the basic functions being perfectly localized in the time-frequency plane. However, this basis is no longer orthonormal. Furthermore, using Daubechies terminology [7], this set of functions is not even a frame, which means that we cannot link the norm of a signal in $\mathbf{L}^2(\mathbf{R})$ to the norm of its coordinates over the basis by a relation such as

$$A\|s\|^2 \leq \sum_{m,n} \|\langle s \mid x_{m,n} \rangle\|^2 \leq B\|s\|^2 \tag{16}$$

where A and B are two positive constants independent of s. This situation is very far from the usual case (for instance with a conventional QAM modulation) where the minimum Euclidian distance between the transmitted signals is equivalent to the minimum Euclidian distance of the code. For digital transmission, where Euclidian distance between signal is a major concern, this is obviously a major drawback.

Coming to this point, it might seem that if we want to construct a modulation scheme making an optimal use of the signal space dimension, based on elementary signals which are correctly localized in both time and frequency domains, then the OFDM approach is more or less a unique solution. However, as will be explained below, such is not the case.

B. The OFDM/OQAM Approach

In standard OFDM, the transmitted data are complex. Each carrier is QPSK modulated (without any filtering), or more generally QAM modulated. Therefore, we shall call it OFDM/QPSK (as in the DAB system [1], [3]) or OFDM/QAM, so that it can be differentiated from the other approaches introduced hereafter.

Now, we consider the situation where the QAM modulation of each carrier is replaced by an Offset QAM (OQAM) scheme. This corresponds to another well known parallel transmission scheme [9]–[11]. In this case also, a link between the modulation scheme and the Hilbertian basis can be established, but for this we need a redefinition of the signal space.

We have defined up to now the signal space as the Hilbert space $\mathbf{L}^2(\mathbf{R})$ with the standard inner product

$$\langle x \mid y \rangle = \int_R x(t) y^*(t) \, dt. \tag{17}$$

Another way to define this Hilbert space is to use another inner product

$$\langle x \mid y \rangle_R = \Re e \int_R x(t) y^*(t) \, dt. \tag{18}$$

Of course, the associated norm is the same in both cases

$$\|x\| = \int_R \|x(t)\|^2 \, dt. \tag{19}$$

A first example of a Hilbertian basis of $\mathbf{L}^2(\mathbf{R})$ with this new inner product is obtained with the prototype function $x(t)$ defined by its Fourier transform

$$X(f) = \begin{cases} \frac{1}{\sqrt{\nu_0}} \cos \pi f / 2\nu_0 & \text{if } |f| \leq \nu_0 \\ 0 & \text{elsewhere.} \end{cases} \tag{20}$$

If we consider $x(t)$ and its Fourier transform, we can notice that $X(f)$ is strictly band-limited and that $x(t)$ decreases as $|t|^{-2}$. This is a much better result than the theoretical limit derived from the Balian–Low–Coifman–Semmes theorem.

Let us define a set of basic functions $x_{m,n}(t)$ obtained by translation in time and frequency of this prototype function

$$x_{m,n}(t) = i^{m+n} e^{2i\pi m \nu_0 t} x(t - n\tau_0)$$

with

$$\nu_0 \tau_0 = 1/2. \tag{21}$$

It can be easily shown [12] that this set of functions is a Hilbertian basis. This result can be extended to any even function whose Fourier transform satisfy the following conditions:

$$\begin{cases} X(f) = 0 & \text{if } |f| \geq \nu_0 \\ |X(f)|^2 + |X^2(f - \nu_0)|^2 = 1/\nu_0 & \text{if } |f| < \nu_0 \end{cases} \tag{22}$$

which corresponds to a half-Nyquist filter.

Then, an OFDM/OQAM signal $s(t)$ can be written as in (6) where the coefficients $a_{m,n}$ take real values representing the transmitted encoded data.

C. The OFDM/MSK Approach

We will define the OFDM/MSK (minimum shift keying) in the same way as an OFDM/OQAM modulation, but using a different prototype function. The name OFDM/MSK has been chosen because each "carrier" can be considered as MSK modulated. The prototype function can be written as

$$x(t) = \begin{cases} \frac{1}{\sqrt{\tau_0}} \cos \pi t / 2\tau_0 & \text{if } |t| \leq \tau_0 \\ 0 & \text{elsewhere.} \end{cases} \tag{23}$$

In fact, this modulation is nothing else than the dual of OQAM. We have just inverted the time and the frequency axis. Therefore, the prototype function is time-limited instead of being band-limited. The main difference is the implementation of the receiver, where the number of tap of the input filter is greatly reduced compared with OFDM/OQAM.

D. The Isotropic Orthogonal Transform Algorithm (IOTA) Approach

All the modulation systems described previously are based on well known signal processing techniques: Short-Time Fourier Transform for OFDM/QAM, para-unitary QMF filter bank for OFDM/OQAM, and Lapped Orthogonal Transform for OFDM/MSK. In all these examples, orthogonality between basic functions is obtained by straightforward methods, making use either of the time or frequency limitation of the prototype function. A different approach is detailed here.

We consider here a Hilbertian basis defined according to an OFDM/OQAM scheme

$$x_{m,n}(t) = i^{m+n} e^{2i\pi m\nu_0 t} x(t - n\tau_0) \tag{24}$$

with

$$\nu_0\tau_0 = 1/2.$$

We shall normalize the time and frequency scales, taking $\nu_0 = \tau_0 = 1/\sqrt{2}$. Let us define $\mathbf{O}$ as the orthogonalization operator which transforms a function x into a function y, where

$$y(u) = \frac{2^{1/4} x(u)}{\sqrt{\sum_k \|x(u - k/\sqrt{2})\|^2}}. \tag{25}$$

The effect of the operator $\mathbf{O}$ is to orthogonalize the function y along the frequency axis.

Let us define the (Woodward) ambiguity function of the function y as

$$A_y(\tau,\nu) = \int y(t + \tau/2) y^*(t - \tau/2) e^{-2i\pi\nu t}\, dt. \tag{26}$$

This function can be seen as a 2D autocorrelation function in the time frequency plane.

The effect of the operator $\mathbf{O}$ can be seen directly on the ambiguity function:

$$A_y(0, m\sqrt{2}) = 0, \quad m \neq 0. \tag{27}$$

Similarly, $\mathbf{F}^{-1}\mathbf{OF}$ orthogonalizes the function y along the time axis ($\mathbf{F}$ being the Fourier transform operator). This can also be seen directly on the ambiguity function

$$A_y(n\sqrt{2}, 0) = 0, \quad n \neq 0. \tag{28}$$

This orthogonalization method is attributed to Henri Poincaré. More recently, it has been used in wavelet theory.

Let us consider now the function $z_\alpha = \mathbf{F}^{-1}\mathbf{OFO}x_\alpha$, with $x_\alpha = (2\alpha)^{1/4} e^{-\pi\alpha u^2}$. Then it can be demonstrated that

$$A_{z_\alpha}(n\sqrt{2}, m\sqrt{2}) = 0, \quad (m,n) \neq (0,0). \tag{29}$$

Therefore, z_α can be taken as the prototype function of a Hilbertian basis. Furthermore, $\mathbf{F}z_\alpha = z_{1/\alpha}$. A special case of particular interest is the function $\Im = z_1$, which is therefore identical to its Fourier transform. This function is represented on Fig. 3. Fig. 4 shows the ambiguity function of $\Im$, compared to the ambiguity function of a gaussian function, which is perfectly isotropic in the time-frequency plane. The top of these two functions are nearly identical, the main difference being at the base of the ambiguity function. We consider the Hilbertian basis $\{\Im_{m,n}\}$ defined as

$$\Im_{m,n}(t) = e^{i\varphi_{m,n}} e^{\sqrt{2}i\pi mt} \Im\left(t - \frac{n}{\sqrt{2}}\right)$$

with

$$\varphi_{m,n} = (m+n)\pi/2. \tag{30}$$

We can alternatively choose

$$\begin{cases} \varphi_{m,n} = 0 & m+n \text{ even} \\ \varphi_{m,n} = \pi/2 & m+n \text{ odd.} \end{cases} \tag{31}$$

Using this Hilbertian basis, we can define a new transform that we shall name IOTA because of the nearly isotropic properties of the prototype function. The IOTA transform of a signal $s(t)$ having a Fourier transform $S(f)$ is defined as

$$\begin{aligned} a_{m,n} &= \Re e \int s(t) \Im^*_{m,n} dt \\ &= \Re e \int s(t) e^{-i\varphi_{m,n}} e^{-\sqrt{2}i\pi mt} \Im\left(t - \frac{n}{\sqrt{2}}\right) dt \\ &= (-1)^{mn} \Re e \int S(f) e^{-i\varphi_{m,n}} e^{\sqrt{2}i\pi nf} \Im\left(f - \frac{m}{\sqrt{2}}\right) df. \end{aligned} \tag{32}$$

The inverse transform being given by

$$\begin{aligned} s(t) &= \sum_{m,n} a_{m,n} \Im_{m,n}(t) \\ &= \sum_{m,n} a_{m,n} e^{i\phi_{m,n}} e^{\sqrt{2}i\pi mt} \Im\left(t - \frac{n}{\sqrt{2}}\right) \\ S(f) &= \sum_{m,n} (-1)^{mn} a_{m,n} e^{i\varphi_{m,n}} e^{-\sqrt{2}i\pi nf} \Im\left(f - \frac{m}{\sqrt{2}}\right). \end{aligned} \tag{33}$$

We use this inverse transform to define an OFDM/IOTA modulation, where the coefficients $a_{m,n}$ of the signal $s(t)$ take real values representing the transmitted encoded data.

E. Performance Analysis of OFDM Schemes

In multipath propagation, the channel can be described by a sum of elementary paths defined by a delay τ, a frequency offset ν, an amplitude ρ and a phase φ. Therefore, the sensitivity of any modulation scheme to multipath propagation can be evaluated from its sensitivity to a static delay and frequency offset. It can be demonstrated that for any OFDM scheme defined by a Hilbertian basis and a

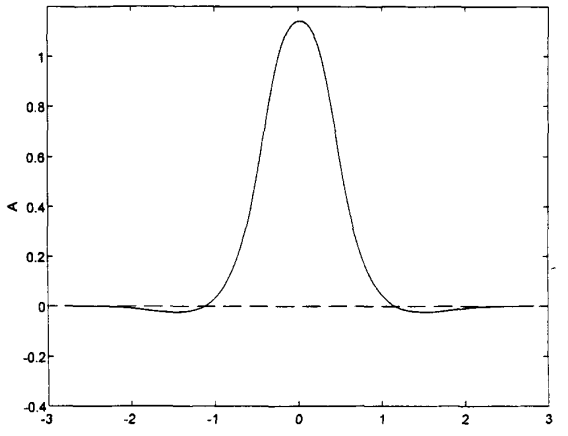


Fig. 3. (a) Linear representation of the IOTA function, (b) logarithmic representation of the IOTA function.

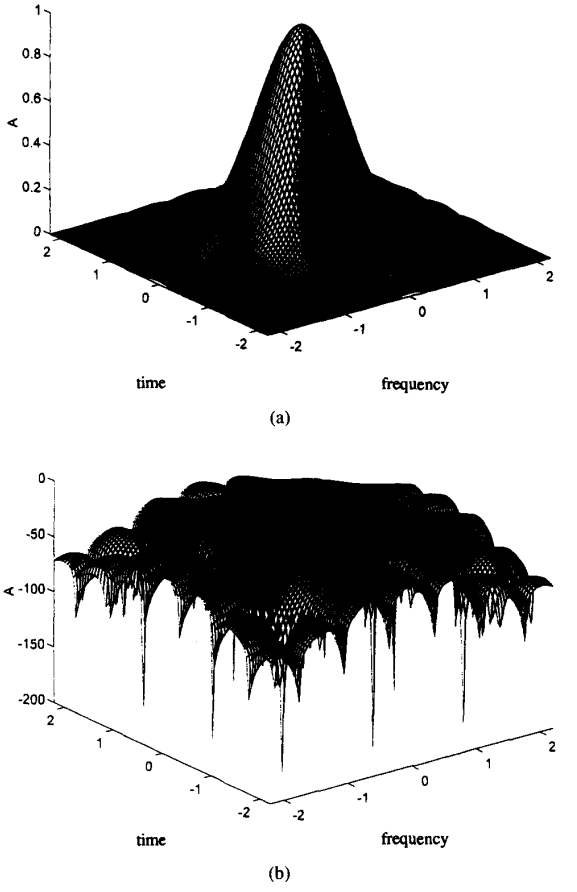


Fig. 4. (a) Linear representation of the IOTA ambiguity function, (b) logarithmic representation of the IOTA ambiguity function.

prototype function $x(t)$, the intersymbol variance can be written as

$$I = (1 - (\Re e[A_x(\tau, \nu)])^2)\sigma^2 \tag{34}$$

where A_x is the ambiguity function of x defined in Section III-D, σ being the variance of the transmitted data. In the case of an even prototype function, the ambiguity function is real, and we can simply write

$$I(\tau, \nu) = (1 - A_x^2(\tau, \nu))\sigma^2. \tag{35}$$

Therefore, the behavior of any OFDM scheme regarding multipath is closely related to the ambiguity function of its prototype function. We can even appreciate this sensitivity with a single parameter by considering the opening of the cone which is tangent to the normalized intersymbol interference function. This opening can be measured by the area of the surface ξ corresponding to the intersection of the cone and the plane of maximum intersymbol interference, i.e., $I(\tau, \nu) = \sigma^2$. Using the Heisenberg inequality, it can be demonstrated that this parameter is always less than unity. If the prototype function is real, this parameter is simply related to its second order moments by the relation:

$$\xi = 1/4\pi\Delta t\Delta f \tag{36}$$

with

$$\begin{cases} \Delta t^2 = \int t^2 \|x(t)\|^2 \, dt \\ \Delta f^2 = \int f^2 \|X(f)\|^2 \, df. \end{cases}$$

Consequently, in the case of OFDM/QAM, $\xi = 0$. If we consider the normalized intersymbol interference function shown in Fig. 5, we find that the sensitivity to a static delay is very high. The reason for this is that Δf is infinite. This does not occur for a frequency shift, because Δt has a finite value.

This high sensitivity to a static delay, and therefore to multipath propagation, seems to be a major drawback of OFDM/QAM. But fortunately there is a way out. As described in Section II, in the case of the OFDM/QAM, we can extend the prototype function with a "guard interval." In such a situation, the system is totally insensitive to the delay as long as this delay is smaller than the guard interval. Fig. 6 shows the effect of the guard interval on

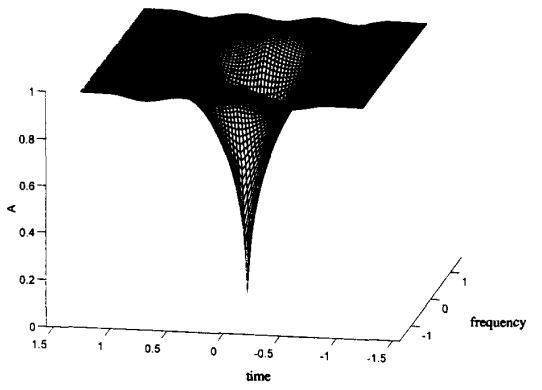


Fig. 5. OFDM/QAM intersymbol interference function without guard interval.

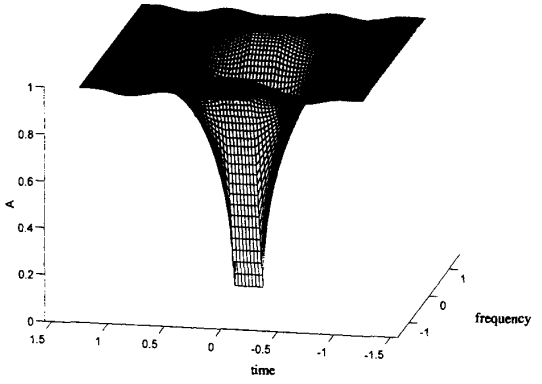


Fig. 6. OFDM/QAM intersymbol interference function in presence of a guard interval.

the intersymbol interference function. Of course, the guard interval does not come for free: first, there is some loss due to the power transmitted during the guard interval, secondly, there is a loss of some space dimensions (i.e., channel capacity). Thus for a given overall spectrum efficiency, we need more bits per dimension, and this leads also to a penalty with respect to the Shannon limit.

The price to be paid greatly depends on the duration of the guard interval which is needed and the spectral efficiency. As an example, if we want an overall spectrum efficiency of 4 b/s/Hz with a guard interval of 1/4 of the useful symbol, we need a coding scheme transmitting 2.5 bits per dimension instead of 2, which represents a 3 dB loss with respect to the Shannon limit. Adding the power wasted in the guard interval itself, the price to pay is 4 dB.

In the case of OFDM/OQAM, there is no guard interval, and such a technique is not even feasible. The parameter ξ equals 0.876. Therefore, OFDM/OQAM is not far from having optimum properties regarding multipath. However, its prototype function decreases relatively slowly in the time domain. This leads to a larger number of taps in the filter of the receiver. This problem is solved with OFDM/MSK, which uses a finite prototype function, while keeping the same ξ parameter. Therefore, OFDM/MSK can be preferred to OFDM/OQAM in most applications.

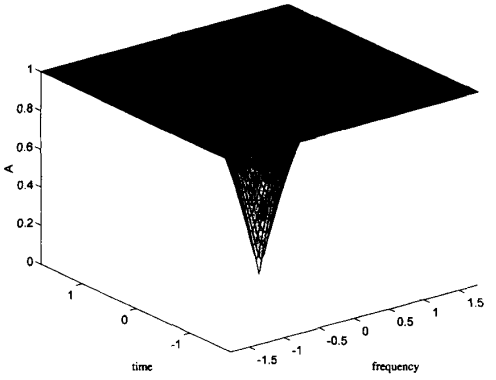


Fig. 7. OFDM/IOTA intersymbol interference function.

The parameter ξ is even improved in OFDM/IOTA and equals to 0.977 in this case. As we know, the theoretical value is 1. This value cannot be reached, for it corresponds to a gaussian prototype, which does not generate an Hilbertian basis. Therefore, it is our conviction (while not yet demonstrated) that OFDM/IOTA is an absolute optimum. The corresponding intersymbol interference function is shown in Fig. 7, and has a nearly circular symmetry.

From the implementation point of view, however, the simplest scheme is the OFDM/QAM: no prefiltering is required. In the case of OFDM/MSK, the prefiltering is quite simple. In the case of OFDM/IOTA, the prototype function has a fast decreasing property (from a mathematical point of view). Therefore, it can be easily truncated. OFDM/OQAM leads to more complex prefiltering, because the prototype function decreases slowly.

Compared with OFDM/QAM: OFDM/OQAM, OFDM/MSK and OFDM/IOTA are more complex in term of FFT by a factor of 2, because of the density of the lattice of the basic functions. Also the channel estimation is more complex. Finally, the choice between these various methods greatly depends on the channel parameters, the spectrum efficiency which is needed and the cost of the receiver. For the time being, the OFDM/QAM scheme is preferred for its simplicity of implementation in hardware (no prefiltering required and simple FFT).

IV. Choice of the Code

A. Introduction

Convolutional channel coding has become widespread in the design of digital transmission systems. One major reason for this is the possibility of achieving real-time decoding without noticeable information losses, thanks to the well known soft-input Viterbi algorithm [13]. Moreover, the same Viterbi decoder may serve for various coding rates by means of puncturing [14], allowing the same silicon product to be used in different applications. Two kinds of convolutional codes are of practical interest: non-systematic convolutional (NSC) and recursive systematic

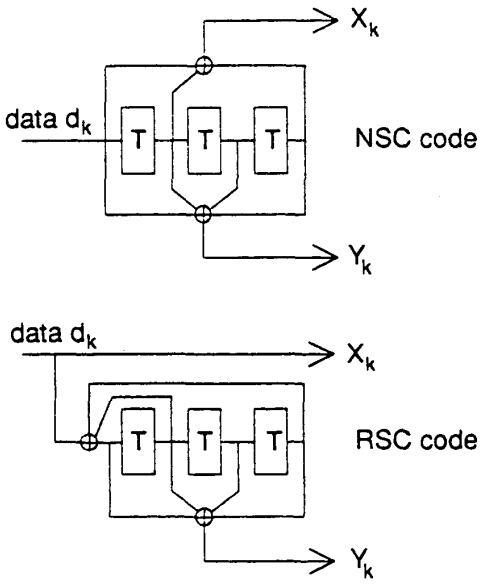


Fig. 8. Nonsystematic convolutional (NSC) and recursive systematic convolutional (RSC) codes with memory code $\nu = 3$ and polynomials 15, 17.

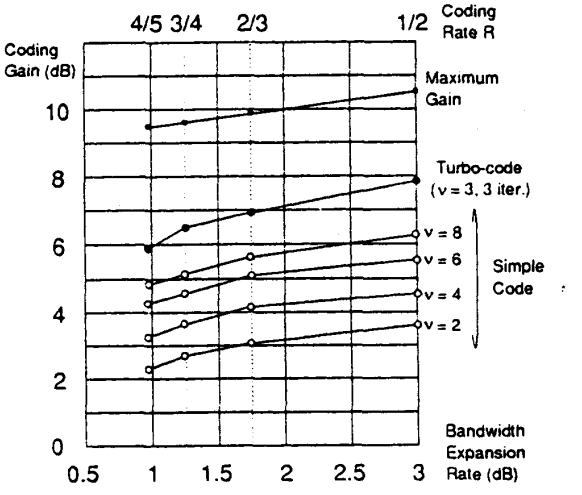


Fig. 9. Coding gains at BER $= 10^{-6}$ of simple NSC codes (3-b quantization [17]) and turbo code (see Section IV-B) with $\nu = 3$, 48 × 32 interleaving and three decoding iterations (4-b quantization), for rates 1/2, 2/3, 3/4, and 4/5 in a Gaussian channel, with QPSK modulation.

convolutional (RSC) codes. Fig. 8 shows both types of codes built with the same parameters: code memory $\nu = 3$ (constraint length $K = \nu + 1 = 4$), polynomials 15, 17, and rate $R = 1/2$. Though RSC codes have the same free distance d_f (i.e., the minimum Hamming distance between two different sequences at the output of the encoder, with the same starting and ending states) as NSC codes and exhibit better performance at low signal to noise ratio (SNR) and/or when punctured [15], only NSC codes have actually been considered for channel coding, except in trellis coded modulations (TCM) [16].

For a given rate, the error-correcting power of convolutional codes, measured as the coding gain at a certain binary error rate in comparison with the uncoded transmission, grows more or less linearly with code memory ν. Fig. 9 (Simple Code [17]) shows the achievable coding gains for different rates, and corresponding bandwidth expansion rates, by using NSC codes with $\nu = 2$, 4, 6, and 8, for a BER of 10^{-6}. For instance, with $R = 1/2$, each unit added to ν gives about 0.5 dB more to the coding gain, up to $\nu = 6$; for $\nu = 8$, the additional gain is lower. Unfortunately, the complexity of the decoder is not a linear function of ν and grows exponentially as $\nu * 2^{\nu}$. The factor 2^{ν} represents the number of states processed by the decoder and the multiplying factor ν accounts for the memory part complexity (metrics and survivor memory). Other technical limitations like the interconnection constraint in the decoder layout make the value of 6 a practical upper limit for ν, for most applications.

In order to obtain high coding gains with moderate decoding complexity, concatenation has proved to be an attractive scheme. Classically, concatenation has consisted in cascading a block code (the outer code, typically a Reed–Solomon code) and a convolutional code (the inner code) in a serial structure. Another—all convolutional—concatenated code, named turbo code [18], [19], has been proposed in the recent past, which offers good performance and reasonable decoding complexity.

B. Turbo Codes

A turbo-encoder is built from a parallel concatenation of two elementary RSC codes, separated by a nonuniform interleaver (defined below). Fig. 10(a) depicts such a structure with $\nu = 3$, leading to a basic $R = 1/3$ encoder. Parallel concatenation, unlike the serial one, enables the whole encoder, and therefore the whole decoder, to run with a single clock. RSC codes have been chosen instead of NSC codes for two reasons. First and mainly, parallel concatenation can only be built with systematic codes. Secondly, for a global coding rate less than or equal to 1/2, both elementary codes have to be punctured and, as mentioned above, RSC codes are superior, especially at low SNR. When a concatenated coding scheme is considered, giving anyhow a steep decrease in the BER curve, it is favorable to have performant codes at low SNR. Nonuniform interleaving consists at once in scattering data, as with usual interleaving, and giving the maximum disorder in the interleaved data sequence. The latter property makes redundancy generation by the two encoders as diverse as possible. The decoder, according to the natural scheme in Fig. 10(b), uses each elementary decoder by feeding back results from one to the other. Data processed by each elementary decoder are

—symbols X_k (or X_n), representing information data, with additive noise,

—its own redundancy Y_{1k} (or Y_{2k}), with additive noise,

—the extrinsic piece of information Z_k (or Z_n), provided by the other decoder.

The output of the global decoder is the output of the elementary decoder DEC1 or DEC2, after desinterleaving, and provides the logarithm of likelihood ratio (LLR) about d_k. The feedback or extrinsic information Z_k (or Z_n) is calculated as the output of the related decoder, reduced by the contributions of inputs X and Z, at the corresponding time, k or n. It provides an estimate of data d_k (or d_n), having a noise weakly correlated with the one which corrupts symbol X_k (or X_n), and the larger the interleaving size, the lower the correlation between these noises. Parameter γ takes into account the noise variance of symbols X (or Y) and Z, which are different.

The natural parallel version of the global decoder can be advantageously replaced by the serial structure of Fig. 10(c), for there is only one piece of extrinsic information to compute and one interleaver. In this case, a second parameter (β) is necessary to compensate for unequal noise variances of X_{2n} and Y_{2k}. Of course, the actual decoder cannot be built as schematically described in Fig. 10(c), because of the different latencies, and the iterative processing that must be carried out. One iteration corresponds to passing through both decoders DEC1 and DEC2.

The basic idea for turbo-decoding relies on this observation: without the feedback loop, decoder DEC2 takes benefit of the whole redundancy information, whereas decoder DEC1 just uses redundancy Y_{1k}; hence, the decoding scheme is not optimal. The role of extrinsic information is then to compensate for this dissymmetry.

Turbo codes with rates higher than 1/3 are easily obtained by puncturing symbols Y_{1k} and/or Y_{2k}. For high rates ($R \geq 5/6$), puncturing X_k may be preferable. The best results achieved with turbo codes have been reported in [18]. By using a very large interleaving matrix (256 × 256) with appropriate nonuniform reading [19], and optimum decoding for $\nu = 4$ elementary codes, a BER of 10^{-5} was obtained with additive white gaussian noise (AWGN) for $R = 1/2$ at $E_b/N_o = 0.7$ dB, after 18 iterations.

In practice, lower sizes for interleaving must be considered, typically 32 × 32 or 48 × 32. Furthermore, optimum decoding has to be replaced by soft-output Viterbi decoding [20] for acceptable complexity. Lastly, the number of iterations for a monolithic integration in silicon is necessarily short: 3 or 4. All this leads to some degradation relatively to the optimum case. Fig. 9 again shows the gain achievable with a monolithic turbo-encoder/decoder VLSI circuit, by comparison with the simple convolutional codes. It is built with two $\nu = 3$ identical codes, a 48 × 32 interleaving matrix and decoding is performed in three iterations. It can be observed by extrapolation that the error-correcting power is comparable to that of a simple code with $\nu = 12$. The material complexity of this circuit being about that of a $\nu = 7$ simple code (including all memories and synchronization functions), the gain in complexity given by the turbo code, relatively to the simple code, is roughly 32.

Because a turbo-decoder is like a pipelined cascade of Viterbi decoders with a short number of states, 8 in the present case, it is suitable to high speed operation. In fact, in such a circuit, the critical path is the memory access, and data rates up to 200 Mb/s in CMOS 0.6 μ technology may be achieved by using parallelism, with one iteration per circuit to avoid dissipation problems.

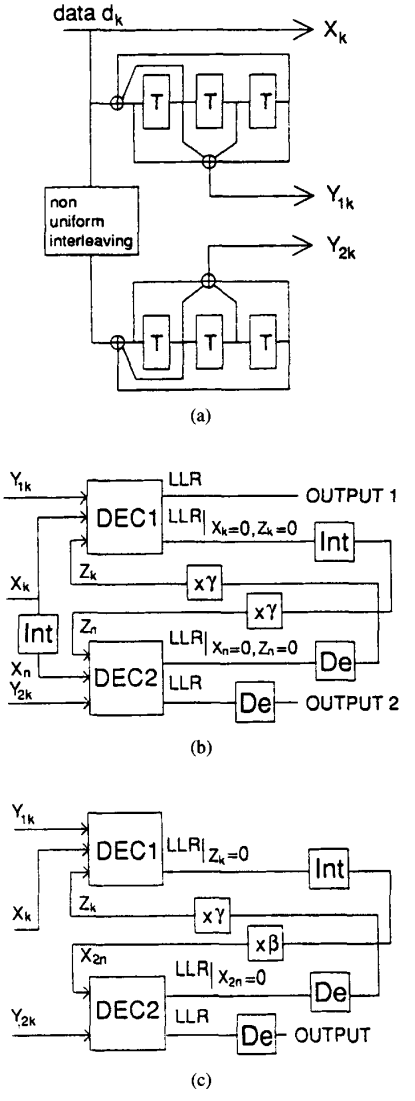


Fig. 10. Turbo-encoder and decoders with natural parallel or serial structure. Int: interleaving, De: desinterleaving. LLR: logarithm of likelihood ratio about d_k. (a) Turbo-encoder, (b) turbo-decoder with parallel structure, and (c) turbo-decoder with serial structure.

C. Synchronization of Turbo-Decoders

Like any concatenated decoder, turbo-decoders have to be node and time synchronized. Because a turbo code is a systematic code, synchronization may directly rely on data frame synchronization, if any. In other cases, internal synchronization without loss of rate, is proposed. It is based on an inverting pattern method and consists in replacing

redundancy Y_{1k} at the encoder output by its inverse at predetermined instants, according to a flag sequence. Decoder DEC1 tries to locate this flag by comparing its binary decision on Y_{1k} with the corresponding input, which is not used by the decoder. All combinations are scanned, which may take some time, depending on the coding rate and the type of modulation chosen. After synchronization, a supervision function in DEC2, using the pseudo-syndrome method [21], is in charge of detecting any *out-of-synchro* situation in order to engage a new synchronization process.

D. Turbo Codes Compared to Classical Concatenation

Since a turbo code is a concatenated code, it is realistic to compare its performance to that of a classical concatenated code. Such a comparison is made with a particular example in Fig. 11. The turbo code is the same as mentioned above: $\nu = 3$, 3 iterations, interleaving 48 $\times$ 32. Its rate is $R = 3/4$. The classical concatenated code uses a NSC inner code with $\nu = 6$ and $R = 4/5$. The outer code is a (204,188) Reed–Solomon (RS) code with 8-b symbols. Its correction capacity is $t = 8$ symbols in the block of 204. The global coding rate of this concatenated code is slightly lower: 0.737. For the latter code, three interleaving depths (the number of blocks of 204 symbols processed by the interleaver) are considered: $I = 1$ (no interleaving between outer and inner codes), 2 and 4. In all cases, the performance is given taking practical limitations into consideration (quantization, truncation length of Viterbi decoders, etc. ...).

For an equivalent material complexity and comparable latencies (5 or 6 kb) of the two encoders/decoders (turbo coding brings no latency), we have to consider the case $I = 2$ for the classical concatenated code and then the performance of the turbo code is better for the practical values generally sought for the BER. As allowed by the principle of turbo codes, the slope of the curve may even be increased by cascading two identical circuits, thus performing six iterations. In this case, the performance of the turbo code is always better, whatever the interleaving depth of the classical concatenated code, but at the expense of a doubled latency.

Since puncturing a turbo-encoder leads to coding rates of the form m/n, m, and $n > m$ being any integers, its association with various modulations is simple especially with respect to clock generation. For instance, this 3/4 rate code may be directly combined with a 16-QAM modulation (see next section) with a modulation rate exactly one third of the data rate. This is not so simple for the classical concatenated code.

E. Turbo Codes Compared to TCM

Thanks to feedback information Z_k which acts as a diversity factor, turbo codes are really robust codes when associated with high spectral efficiency modulations and/or when used on fading channels. Particularly, it has been shown [22] that turbo codes offer much better protection than TCM on both Gaussian and Rayleigh channels. Fig. 12 gives the example of a 3 b/Hz/s transmission on a

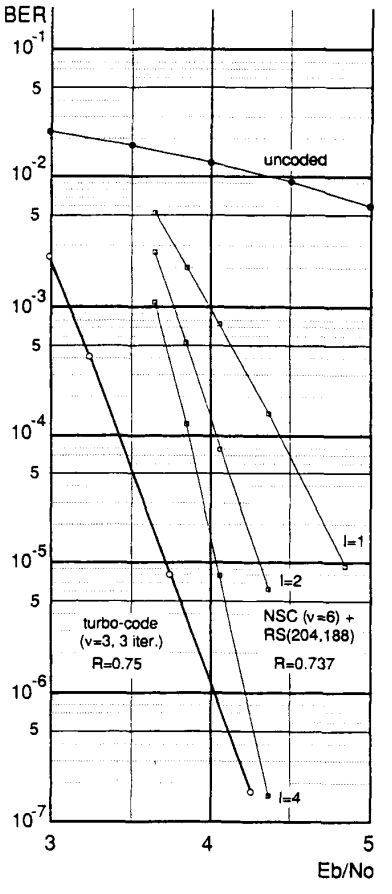


Fig. 11. Comparison of a turbo code (same as Fig. 9, $R = 0.75$) and a classical serial concatenated code ($R = 0.737$) performance in a Gaussian channel (I: interleaving depth).

Gaussian channel. Curve (a) is for a 16 QAM associated with the same turbo code as presented above ($\nu = 3$, 3 iterations, interleaving 48 $\times$ 32, $R = 3/4$, integrated circuit limitations). All information bits are protected and decoding is carried out in a similar way as the pragmatic approach proposed by Viterbi *et al.* [23]. Indeed, the same turbo-encoder/decoder, with puncturing, may be used in association with various modulations and different spectral efficiencies. It just needs specific transcoding between the encoder and the modulator, and between the demodulator and the decoder, to realize adaptation from binary symbols to nonbinary symbols, and vice-versa. Curve (b) gives the theoretical performance of the 64-state TC-16QAM, which is less favorable for BER lower than 10^{-3}, despite a greater decoding complexity. Another, still more complex possible transmitting scheme (curve (c)) associates the 64-state TC-16QAM and the RS (204,188) code with interleaving depth $I = 2$. The spectral efficiency is slightly lower (2.76 b/Hz/s). The theoretical performance is comparable to the turbo code solution. These results may be extended to various comparisons that can be made between TCM and noncoded modulations with channel turbo-coding, whatever

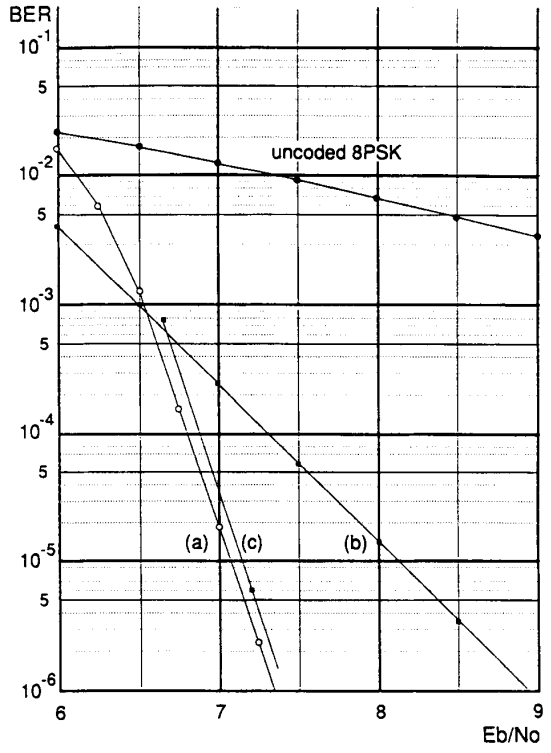


Fig. 12. Comparison of a turbo code and a TCM performance in a Gaussian channel. Curve (a): 16 QAM with $R = 3/4$ turbo code (same as Fig. 9), spectral efficiency: 3 b/Hz/s. Curve (b): 64-state TC-16QAM (theoretical), spectral efficiency: 3 b/Hz/s. Curve (c): 64-state TC-16QAM with (204,188) RS code ($I = 2$), spectral efficiency: 2.76 b/Hz/s.

the spectral efficiency. In a Rayleigh channel, turbo codes may outclass TCM by more than 8 dB [22]. This is mainly because all bits are protected by the code, which moreover offers a good performance despite puncturing.

Thus turbo codes represent a suitable solution for channel coding in most digital communication systems. With cost-effective VLSI circuits, good performance (and often the best at the present time) can be obtained for various spectral efficiencies. For digital terrestrial television, a specific cascadable VLSI turbo-decoder has been designed with 16-state soft-output Viterbi decoders and a 64 × 32 interleaving matrix. The performance it offers is given in the following section.

V. System Performances

The COFDM system has the potential to use an echo to improve its performance. This has often been shown in demonstrations by adding a 0 dB echo with 20 μs delay for example to a transmission in white Gaussian noise which already includes a direct path and another 0 dB echo with 15 μs delay. The power of the second echo (the 20 μs one) is used by the COFDM system to reduce the bit error rate (BER)! How is it possible?

Any digital transmission system in presence of linear channel distortions (nonlinear distortion usually caused by equipment limitation is not considered here) shows performances which depend on two factors:

—the carrier-to-noise ratio (C/N)

—the characteristics of the distortions.

When the channel is perfect (this is an unrealistic situation), the BER performance is referred to as the Gaussian channel performance. When the channel is strongly distorted, including echoes that can be as powerful as the direct path, the BER performance is referred to as the frequency selective channel (worst case) performance.

In the COFDM system, the difference between these two BER performances depends on the strength of the code used.

The principle of additivity of the echoes can be interpreted in the following way (see Fig. 13): starting from the Gaussian channel (only one direct path), with a BER of 10^3, the addition of an echo as strong as the direct path will lead to the following two consequences:

—a 3 dB increase of the C/N ratio,

—a change from the Gaussian performance curve to the Frequency Selective performance curve.

If the code used provides a distance between the two curves lower than 3 dB, then the BER will be decreasing when this first echo is added; this is the case in the Eureka 147 DAB system.

If the code used provides a distance between the two curves higher than 3 dB (this may be the case in the TV broadcasting application where the need for a high spectrum efficiency makes the coding performance demand more challenging), the BER will then be increased by this first "0 dB echo."

When afterwards further echoes are added, they will all contribute to increasing the C/N ratio. However, the performance curve will remain at worse the frequency selective one, and therefore the BER will decrease with the insertion of further echoes. In that sense, COFDM benefits from the power addition of echoes to improve performance, as was observed in laboratory measurements using real time hardware.

The COFDM system will operate under the above mentioned condition as long as the multipath delay is contained

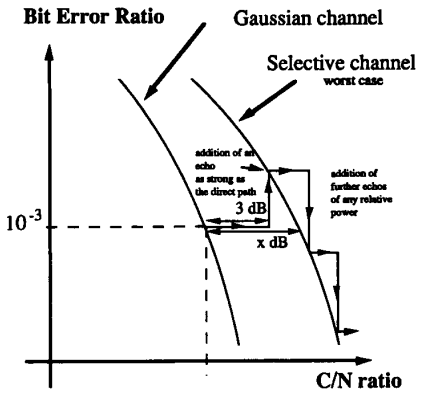


Fig. 13. Illustration of the principle of additivity of the echoes.

within the extent of the guard interval. However, an echo exceeding the guard interval duration will not lead to an abrupt degradation of the performances: it will remain graceful. An echo of power W delayed by τ will present a constructive component C and an interfering component I. It can be shown that the powers of these two components are given by

$$C = W \cdot f(\tau) \quad \text{and} \quad I = W \cdot (1 - f(\tau)) \tag{37}$$

where

$$f(\tau) = 1 \text{ when } \tau \leq \Delta, \text{ (no interfering component)}$$

$$f(\tau) = \left(1 - \frac{\tau - \Delta}{T_0 - \Delta}\right)^2, \text{ when } \Delta \leq \tau \leq T_0$$

and

$$f(\tau) = 0 \text{ when } \tau > T_0$$

where T_0 represents the total length of the COFDM symbol, and Δ represents the duration of the guard interval, as presented in Section II. In presence of a received signal constituted by i paths of powers W_i, with individual constructive component C_i and interfering component I_i, the global C/I ratio then equals

$$\frac{C}{I} = \frac{\sum_i C_i}{\sum_i I_i}. \tag{38}$$

This formula allows the computation of the C/I ratio in the case where echoes are longer than the guard interval. Relating it to the Bit Error Rate versus C/N curve permits to quantify the effect of echoes with very long excess delay on the system performance, as it can be shown that the interfering component has the characteristic of a Gaussian noise.

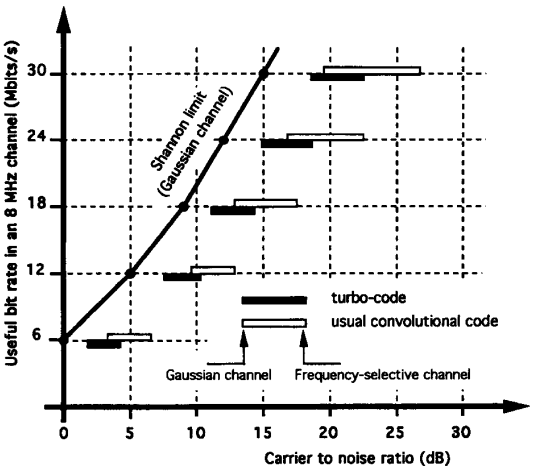
Fig. 14 gives the performance of the COFDM system in terms of possible tradeoffs between the available bit rate in an 8 MHz channel, and the carrier to noise ratio necessary to obtain an unimpaired picture quality. These results include the overheads due to the guard interval, the frequency guard bands and the references for receiver synchronization, so that the effective spectrum efficiency is reduced by a factor 3/4. The performances are given for two types of codes: usual convolutional codes ($\nu = 6$) and turbo codes ($\nu = 4$, 3 iterations).

The Shannon limit is the theoretical limit for error-free digital transmission over a Gaussian channel (channel only affected by additive white Gaussian noise).

The figure has to be understood in the following way: the left end of the segment gives the performance of the coded modulation in the Gaussian channel; the right end of the segment gives the performance in the most difficult channel (the frequency-selective channel), in which full notches due to echoes as strong as the main signal may appear.

VI. COFDM versus Linear Equalization of Coded Single Carrier Schemes

Alternative solutions to the COFDM technique are known as single carrier modulation techniques. The main problem



Useful bit rate (Mbits/s)	Code rate	Modulation associated to each carrier
6	1/2	4-PSK
12	2/4	16-QAM
18	3/4	16-QAM
24	4/6	64-QAM
30	5/6	64-QAM

Fig. 14. COFDM performances.

of single carrier modulations is the handling of echoes and the resulting intersymbol interference (ISI): unlike COFDM, the symbol length is very small (typically 0.15 μs), and echoes may occur, that are much longer than a symbol (i.e., up to tens of μs in an urban area, up to more than 100 μs in a Single Frequency Network). This means that the received signal samples are actually a combination of many consecutive symbols. The coefficients of this combination reflect the channel impulse response. The comparison is limited to linear equalization, which is often considered for long term echoes mitigation. However, it is known that techniques such as Decision Feedback Equalization are also considered for their efficiency in the presence of strong echoes.

Linear equalization roughly consists in inverting this combination by combining several consecutive received samples with the proper coefficients. Since the coefficients depend on the channel response, such channel response must hence be evaluated. This may be achieved through a training sequence (start-up), followed by a blind adaptation of the coefficients (tracking). From a theoretical point of view, the performance of ideal linear equalization can be assessed by considering frequency domain equalization, where the transfer function of the equalizer is given by the formula

$$G(\nu) = \frac{H^*(\nu)}{\|H(\nu)\|^2 + \sigma_n^2/\sigma_d^2} \tag{39}$$

where $H(\nu)$ is the channel transfer function, and σ_n^2 and σ_d^2 represent the variance of the noise and the data symbols, respectively.

This equalizer is known to achieve the Minimum Mean Square Error criterion, thus minimizing the combined effect

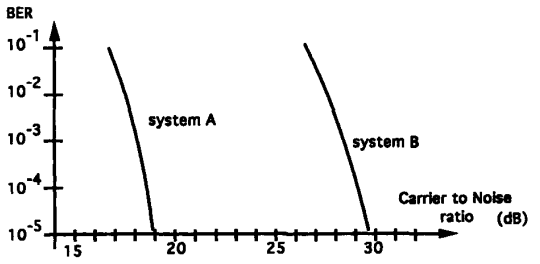


Fig. 15. Compared performances of COFDM and coded single carrier scheme under channel profile $P0$.

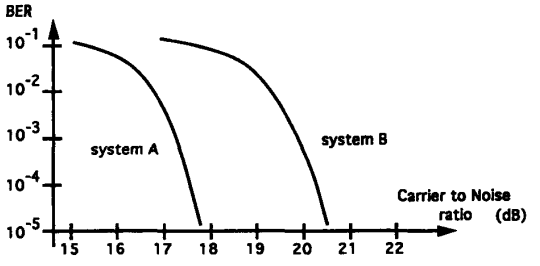


Fig. 16. Compared performances of COFDM and coded single carrier scheme under channel profile $P3$.

of additive noise and residual intersymbol interference. In order to compare the performances of such equalization applied to a single carrier modulation scheme, with those of COFDM, the following hypotheses are retained

—system A is a COFDM system using a 64-QAM rate 2/3 turbo coded ($\nu = 4$, 3 iterations) modulation on each carrier; the receiver applies a coherent demodulation with an assumed perfect knowledge of the transmission channel.

—system B is a single carrier scheme using the same 64-QAM rate 2/3 turbo coded modulation; the receiver applies the linear equalization defined by the above transfer function $G(\nu)$ (that assumes also perfect knowledge of the transmission channel), and coherent demodulation is applied.

These two systems are passed through the same transmission channel, which is defined by the following impulse response, containing one direct path and a single echo

$$h(t) = \delta(t) + \alpha \cdot \delta(t - \tau). \tag{40}$$

The parameters τ is chosen to be much larger than the inverse of the signal bandwidth (about 100 times), so that the simulation results are independent of the echo phase. Three values of the parameter α have been retained: 1, $1/\sqrt{2}$, 1/2. This corresponds to an echo attenuation of 0 dB, 3 dB, and 6 dB, respectively, and the three channel profiles are denoted by $P0$, $P3$ and $P6$. The results obtained are plotted in Figs. 15–17.

The following conclusions can be derived from these results:

—in the case of COFDM, the performance improvement between $P0$ and $P3$, as well as between $P3$ and $P6$, is approximately 1 dB, which confirms that the performance dispersion of turbo codes between a "Gaussian-like" and a "Rayleigh-like" channel is rather limited; this is a good indication of the intrinsic code performance.

—despite the use of the same code associated to the same modulation in the case of optimal linear equalization of the single carrier scheme, the performance improvement between $P0$ and $P3$ is of the order of 10 dB, and still a little more than 2 dB improvement is achieved with profile $P6$. This shows that the statistics of the maximum likelihood decoder inputs are of utmost importance and strongly condition the decoding performance. In the case of COFDM, the maximum likelihood decoder is able to benefit from full channel knowledge thanks to its weighted inputs (i.e., metrics of the Viterbi decoder) that represent a reliable image of the channel distortions.

—the performance differences between systems A and B, at 10^{-5} BER, are equal to 11 dB, 2.5 dB, and 1 dB for channel profiles $P0$, $P3$, and $P6$, respectively.

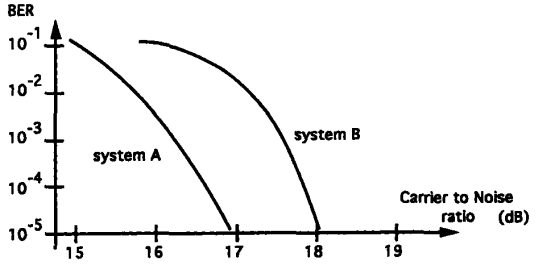


Fig. 17. Compared performances of COFDM and coded single carrier scheme under channel profile $P6$.

VII. Broadcasting Network Architecture

The following paragraph examines COFDM in terms of a "broadcasting system" and shows that the new horizons opened up by this technique are even more attractive when looking at the implementation of a broadcasting system as a whole.

A. Gap-Fillers

Since the system has been designed to take advantage of echoes, it is possible to operate the system in the presence of echoes which are created voluntarily. The initial application of this idea provides the means of eliminating residual shadowed areas by using passive reflectors or small active relays without having to change the carrier frequency for the reflected or relayed signal. The signal is picked up at a location where reception conditions are satisfactory. It is then reamplified and rebroadcast at the same frequency toward the shadowed area that is to be suppressed. These "gap-fillers" do not require any additional frequency and their intrinsic simplicity ensures very low costs. If we consider that the total number of relays necessary to provide national TV coverage in a country such as France adds up to several thousand units, it is easy to see the benefits of

such an approach. Since 1988, full-scale tests around 900 MHz have been carried out in Rennes, France, and have successfully shown the feasibility of these "gap-fillers."

B. Dense Networks and Power Efficiency of the System

If one imagines a network of terrestrial transmitters distributed over a given territory, all of them time-synchronized and broadcasting the same signal on the same frequency, then the useful power received at the input of the receiver is the sum of the incoming power from each transmitter. The various incoming signals are seen as echoes of the same signal and combine positively if their temporal spread is compatible with the selected duration of the guard interval. In other words, COFDM enables constructive overlapping of the various areas of transmitter coverage. The linking, by this technique, of the largest possible number of interdependently operating transmitters offers many advantages:

- —The broadcasting infrastructure is less expensive since it avoids the need for excessively powerful transmitters whose cost increases in relation to the area to be covered.
- —It enables the launch of a new service with minimum initial investment, then a gradual expansion and improvement in the area served.
- —It makes more efficient use of the power transmitted as shown qualitatively in the diagram below (Fig. 18).
- —Unlike analog systems which cater for extremely variable signal to noise ratios, digital systems are very sensitive to a threshold effect within which there is marked quality loss and beyond which the quality is no longer improved. By enabling a more accurate designation of the area of coverage, and greater uniformity with respect to the received power, dense networks can considerably reduce this drawback, which is specific to digital systems. Indeed, the concept could be summarized as: "putting the right power at the right place."
- —It improves control of co-channel interference at the edge of the coverage area, and dramatically reduces the frequency reuse distances. This is a possible solution to the problem of spectrum sharing in dense urban areas.
- —In addition to the time and frequency diversities provided by COFDM, it brings spatial diversity into transmission.

An experimental network comprising two transmitters operating at 60 MHz, which was brought into service around the town of Rennes, confirmed the fact that the overall coverage area served by two transmitters operating simultaneously is significantly greater than the sum of the coverage areas for each of the transmitters operating independently.

C. Single Frequency Networks and Spectral Efficiency of the System

The concept of a network of transmitters operating synchronously on the same frequency conflicts with conventional techniques which require the broadcasting of the same program from geographically adjacent transmitters on two distinct frequencies. If we consider that the 45 8-MHz channels in the UHF band allocated to television in France are totally saturated by five national channels, COFDM can provide the means of increasing this number nine fold. If we also consider a rate of 24 Mb/s in each channel, then the combination of COFDM and source coding provides a gain in terms of spectral occupation in a ratio of $9 \times 4 = 36$, assuming that 6 Mb/s can provide a standard quality picture. Fig. 19 shows the dramatic overall spectral efficiency achieved by a system based on COFDM.

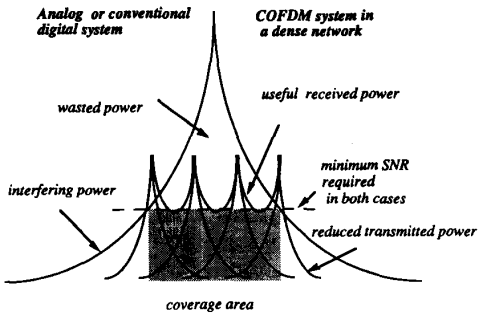


Fig. 18. Power efficiency of dense networks.

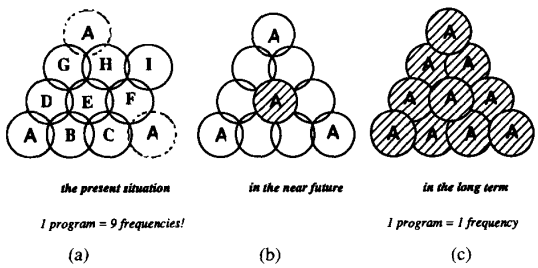


Fig. 19. (a) Frequency planning for conventional analog or digital systems, (b) introduction of a digital service in an analog context, and (c) frequency efficiency of a single-frequency network.

VIII. Conclusion

The tremendous improvements of the integrated circuits technology today allows the real-time implementation of communication systems which could hardly have been simulated with the most powerful computers one decade ago. This is the case of the COFDM system, which relies on the conjunction of very basic principles of the communication theory.

In order to jointly resolve the problems of multipath propagation in terrestrial broadcasting and of optimization of the spectrum usage through Single Frequency Networks, this technique uses many tools which are now common to the 1990's engineer: Time-frequency transforms, maximum likelihood Viterbi decoding, multidimensional filtering. The result obtained is not only an efficient modulation scheme matched to high bit-rate transmission in a dispersive channel, it also offers a promising approach to terrestrial broadcasting networks. Today, there are good reasons to believe

that the COFDM technology will be used to ensure the migration from analog to digital networks, for radio first, and for television afterwards.

The ruggedness of the system has been verified by real-time implementation of a TV COFDM modem offering a bit-rate of 21 Mb/s and able to work in a selective channel affected by echoes as strong as the main path, with 1 dB implementation margin compared to the theoretical result.

Depending on the time pressure of the standardization process, as well as on the tradeoff between performance and receiver cost, different options of implementation can be selected, ranging from the classical implementation of COFDM to the most sophisticated design described in this article under the name of IOTA.

REFERENCES

[1] C. Dosch, P. Ratliff, and D. Pommier, "First public demonstrations of COFDM/MASCAM: A milestone for the future of radio broadcasting," EBU Rev.—Tech. no. 232, Dec. 1988.

[2] D. C. Cox and R. P. Leck, "Distributions of multipath delay spread and average excess delay for 910 MHz urban mobile radio paths," *IEEE Trans. Vehic. Tech.*, vol. VT-26, Nov. 1977.

[3] D. Pommier and Yi. Wu, "Interleaving or spectrum-spreading in digital radio intended for vehicles," EBU Rev.—Tech. no. 217, June 1986.

[4] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency division multiplexing using the discrete Fourier transform," *IEEE Trans. Commun. Tech.*, vol. COM–19, Oct. 1971.

[5] J. F. Hélard and B. Le. Floch, "Trellis coded orthogonal frequency division multiplex for digital video transmission," *Proc. Globecom 91*.

[6] E. Jensen, T. Hoholdt, and J. Justesen, "Double series representation of bounded signals," *IEEE Trans. on Inf. Theory*, vol. 34, July 1988.

[7] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inform. Theory*, vol. 36, Sept. 1990.

[8] D. Gabor, "Theory of communication," *J. Inst. Elect. Eng.*, London, vol. 93, no. 3, pp. 429–457, 1946.

[9] R. W. Chang, "Synthesis of band-limited orthogonal signals for multichannel data transmission," *Bell Syst. Tech. J.*, vol. 45, pp. 1775–1796, Dec. 1986.

[10] B. R. Saltzberg, "Performance of an efficient parallel data transmission system," *IEEE Trans. Commun. Tech.*, vol. COM-15, pp. 805–811, Dec. 1967.

[11] B. Hirosaki, "An analysis of automatic equalizers for orthogonal multiplexed QAM systems," *IEEE Trans. Commun.*, vol. COM-28, Jan. 1980.

[12] ——, "A maximum likelihood receiver for an orthogonally multiplexed QAM system," *Proc. IEEE Int. Conf. on Commun.*, vol. 1, 1984.

[13] G. D. Forney, "The Viterbi algorithm," *Proc. IEEE*, vol. 61, pp. 268–278, Mar. 1973.

[14] J. B. Cain, G. C. Clark, and J. M. Geist, "Punctured convolutional codes of rate $(n-1)/n$ and simplified maximum likelihood decoding," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 97–100, Jan. 1979.

[15] P. Thitimajshima, "Les codes convolutifs récursifs systématiques et leur concaténation," dissertation, l'Université de Bretagne Occidentale, no. 284, Brest, France, 1993.

[16] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 55–67, Jan. 1982.

[17] Y. Yasuda, K. Kashiki, and Y. Hirata, "High-rate punctured convolutional codes for soft-decision Viterbi decoding," *IEEE Trans. Commun.*, vol. COM-32, Mar. 1984.

[18] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon error-correcting coding and decoding: Turbo-codes," *Proc. IEEE Int. Conf. on Commun. (ICC '93)*, pp. 1064–1070, Geneva, May 1993.

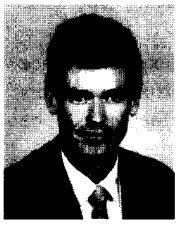
[19] C. Berrou and A. Glavieux, "Turbo-codes: General principles and applications," *Proc. 6th Int. Tirrenia Workshop on Digital Commun.*, pp. 215–226, Tirrenia, Italy, Sept. 1993.

[20] C. Berrou, P. Adde, E. Angui, and S. Faudeil, "A low complexity soft-output Viterbi decoder architecture," *Proc. IEEE Int. Conf. on Commun. (ICC '93)*, pp. 737–740, Geneva, May 1993.

[21] C. Berrou and C. Douillard, "Pseudo-syndrome method for supervising Viterbi decoders at any coding rate," *Electron. Lett.*, vol. 30, no. 13, June 1994.

[22] S. Le Goff, A. Glavieux, and C. Berrou, "Turbo-codes and high spectral efficiency modulation," *Proc. IEEE Int. Conf. on Commun. (ICC '94)*, pp. 645–649, New Orleans, May 1994.

[23] A. J. Viterbi, E. Zehavi, R. Padovani, and J. K. Wolf, "A pragmatic approach to trellis-coded modulation," *IEEE Commun. Mag.*, vol. 27, no. 7, pp. 11–19, July 1989.



Bernard Le Floch was born in Quimper, France, in 1958.

He joined the CCETT (France Telecom and Télédiffusion de France joint research center for telecommunications and broadcasting) in 1985. In 1987 he worked on digital audio broadcasting developments within the EUREKA 147 DAB project, and in 1988 became Head of the Digital Modulation and Channel Coding Department. Since 1990 his work has focused on the development of digital terrestrial television, and has been participating in the European RACE dTTb project as chairman of the Channel Coding and Modulation group. He is now Head of the CCETT division in charge of cable, satellite, and terrestrial digital broadcasting.

Michel Alard (Member, IEEE) was born in Plouénan, France, in 1951. He graduated from the Ecole Polytechnique, Paris, in 1976 and the Ecole Nationale Supérieure des Telecommunications, Paris, in 1978.

From 1979 to 1982 he was with the Société Francaise de Production and Télédiffusion de France, where he was involved with TV broadcasting. From 1982 to 1988 he was with the CCETT, where his work focused on satellite TV broadcasting and digital audio broadcasting. He has developed and promoted the COFDM approach as the modulation and channel coding scheme for the DAB EUREKA 147 project, for which he was issued several patents. In 1988 he joined Matra Communication, and was engaged in the development of the GSM mobile communication system, either on base station and mobile phones. In 1993 he co-founded WAVECOM, Paris, of which he is now President. His current interest is the development and licensing of new technologies for mobile communication.

Claude Berrou (Member, IEEE) was born in Penmarc'h, France, in 1951. He received the electrical engineering degree from the Institut National Polytechnique, Grenoble, France, in 1975.

In 1978 he joined the Ecole Nationale Supérieure des Telecommunciations de Bretagne (France Telecom University), where he is currently a Professor of Electronic Engineering. His research focuses on joint algorithms and VLSI implementations for digital communications, especially error-correcting codecs and synchronization techniques.