Improving the Spectral Efficiency of FDM-CPM Systems through Packing and Multiuser Processing

Amina Piemontese, Nicolò Mazzali, and Giulio Colavolpe

Università di Parma, Dipartimento di Ingegneria dell'Informazione, Viale G. P. Usberti 181/A, Parma - ITALY

Abstract— In frequency division multiplexed (FDM) systems, spectral efficiency can be increased by reducing the spacing between two adjacent channels, thus increasing the relevant interference and possibly accounting for it at the receiver. In this paper, we consider a FDM system where each user employs a continuous phase modulation (CPM), serially concatenated with an outer code through an interleaver, and iterative detection/decoding. We show that, by taking into account the increased interference using properly designed multiuser detection and synchronization schemes, it is possible to implement transmission schemes with unprecedented spectral efficiency at a price of a limited complexity increase with respect to a classical single-user receiver which neglects the interference.

I. INTRODUCTION

Spectral efficiency of frequency division multiplexed (FDM) systems can be increased by reducing the spacing between two adjacent channels, thus allowing overlap in frequency and hence admitting a certain amount of interference [1], [2]. This aspect has been investigated from an information-theoretic point of view for linear [3] as well as continuous phase modulations (CPMs) [4], showing that a significant improvement can be obtained through packing even when at the receiver side a single-user detector is employed. When a multiuser receiver is adopted, the benefits in terms of spectral efficiency can be even larger and the signals can be packed denser and denser [1]–[4].

Since, as known, the complexity of the optimal multiuser detector increases exponentially with the number of channels, suboptimal detection schemes are required. In the case of a satellite FDM system using linear modulations, the adoption of reduced-complexity multiuser detection (MUD) algorithms borrowed from the literature on code division multiple access (CDMA) is investigated in [1]–[3] showing that these techniques work well also in this scenario. Although this is, in principle, possible for CPM systems as well, a new reduced-complexity MUD algorithm for additive white Gaussian noise (AWGN) channel is derived in [5] (see also [6]) based on factor graphs (FGs) and the sum-product algorithm (SPA) [7]. This latter framework, often used in the past to reinterpret known algorithms, is very useful for deriving new detection schemes with an unprecedented complexity/performance trade-off [8] or for applications where traditional probabilistic methods fail [9]. In this case, the new algorithm designed in [5], [6] by using this framework outperforms all other suboptimal MUD algorithms both from performance and complexity point of view.

But a denser packing has an impact not only on the detection algorithm. In fact, once satisfactorily suboptimal MUD algorithms are available, other subsystems become critical. In particular, carrier synchronization schemes able to cope with the increased interference must be adopted.

In this paper, we will focus on CPM signals, since they are often employed in satellite communications for their robustness to non-linearities, stemming from the constant envelope, their claimed power and spectral efficiency, and their recursive nature which allows to employ them in serially concatenated schemes [10], [11]. In particular, we will consider CPMs serially concatenated with an outer code through an interleaver, and iterative detection/decoding. After the system model description of Section II, an information-theoretic analysis aimed at finding the optimal spacing among adjacent channels is described in Section III. The suboptimal reduced-complexity MUD scheme proposed in [5], [6] is reviewed in Section IV. This scheme is obtained by means of some graphical manipulations on the FG representing the joint a posteriori probability mass function (pmf) of the transmitted symbols, rather than assuming the interference as Gaussian, as in schemes from CDMA literature [12]–[14]. Synchronization issues are discussed in Section V. The performance analysis is reported in Section VI and, finally, some conclusions are drawn in Section VII.

II. SYSTEM MODEL

We assume that the channel is shared by U independent users. Without loss of generality, we consider synchronous users, all employing the same modulation format and transmitting at the same power, and an AWGN channel. The extensions to the cases of asynchronous users and of a channel with an unknown and possibly time-varying phase are discussed later. The case of users transmitting at a different power is considered in [5].

Each user transmits N symbols and we denote by $\alpha_n^{(u)}$ the symbol transmitted by user u at discrete-time n, which takes on values in the M-ary alphabet $\{\pm 1, \pm 3 \cdots \pm (M-1)\}$. Moreover, $\boldsymbol{\alpha}^{(u)} = (\alpha_0^{(u)}, \ldots, \alpha_{N-1}^{(u)})^T$ is the vector of the N symbols transmitted by user u and we also denote $\boldsymbol{\alpha}_n = (\alpha_n^{(1)}, \ldots, \alpha_n^{(U)})^T$ and $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_0^T, \ldots, \boldsymbol{\alpha}_{N-1}^T)^T$.[1] The complex envelope of the received signal can be written as

$$r(t) = \sum_{u=1}^{U} s^{(u)}(t, \boldsymbol{\alpha}^{(u)}) \exp\{\jmath 2\pi f^{(u)} t\} + w(t) \qquad (1)$$

where $w(t)$ is a zero-mean circularly symmetric white Gaussian noise process with power spectral density $2N_0$ (N_0 is assumed perfectly known at the receiver), $f^{(u)}$ is the difference between the carrier frequency of user u and the frequency assumed as reference for the computation of the complex envelope, and $s^{(u)}(t, \boldsymbol{\alpha}^{(u)})$ is the CPM information-bearing

---

[1]In the following, $(\cdot)^T$ denotes transpose and $(\cdot)^H$ transpose conjugate.

signal of user u which reads

$$s^{(u)}(t, \boldsymbol{\alpha}^{(u)}) = \sqrt{\frac{2E_S}{T}} \exp\left\{ j2\pi h \sum_{n=0}^{N-1} \alpha_n^{(u)} q(t - nT) \right\}. \quad (2)$$

In (2), E_S is the energy per information symbol of user u, T the symbol interval, $q(t)$ the *phase-smoothing response*, and $h = r/p$ the modulation index (r and p are relatively prime integers). The derivative of the function $q(t)$ is the so-called *frequency pulse* of length L symbol intervals. In the generic time interval $[nT, nT + T)$, the CPM signal of user u is completely defined by symbol $\alpha_n^{(u)}$ and state $\sigma_n^{(u)} = (\omega_n^{(u)}, \phi_n^{(u)})$ [15], where

$$\omega_n^{(u)} = (\alpha_{n-1}^{(u)}, \alpha_{n-2}^{(u)}, \ldots, \alpha_{n-L+1}^{(u)}) \quad (3)$$

is the correlative state and $\phi_n^{(u)}$ is the phase state which can be recursively defined as

$$\phi_n^{(u)} = [\phi_{n-1}^{(u)} + \pi h \alpha_{n-L}^{(u)}]_{2\pi} \quad (4)$$

where $[\cdot]_{2\pi}$ denotes the "modulo 2π" operator, and takes on p values. In the following, we define $\boldsymbol{\sigma}_n = (\sigma_n^{(1)}, \ldots, \sigma_n^{(U)})^T$ and $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_0^T, \ldots, \boldsymbol{\sigma}_N^T)^T$.

Considering the useful component of the received signal (1), in the time interval $[nT, nT + T)$ only $(pM^L)^U$ possible waveforms are allowed. Hence, a set of sufficient statistics for detection can be obtained through projection of each slice of duration T of the received signal $r(t)$ onto an orthonormal basis, of cardinality at most M^{LU}, of the signal space of these possible waveforms. The use of an orthonormal basis will ensure that the noise components are independent and identically distributed (i.i.d.) complex Gaussian random variables with independent components, each with mean zero and variance N_0. In practical receivers, an approximated set of sufficient statistics is obtained through the technique described in [16]. It is assumed that the useful signal component in $r(t)$ is band-limited—although this is not strictly true in the case of CPM signals, whose spectrum has an infinite support—with bandwidth lower than $\eta/2T$, where η is a proper integer. The approximated statistics can be obtained by extracting η samples per symbol interval from the received signal prefiltered by means of an analog low-pass filter which leaves unmodified the useful signal and has a vestigial symmetry around $\eta/2T$.[2] The condition on the vestigial symmetry of the analog prefilter ensures that the noise samples are i.i.d. complex Gaussian random variables with independent components, each with mean zero and variance $\Xi^2 = N_0 \eta/T$.

Although not necessary in the derivation of the algorithms, since it applies unmodified independently of the employed orthogonal basis, the only difference being the noise variance, in the following we will assume a sufficient statistic obtained through oversampling. We will denote by $r_{n,m}$ the m-th received sample ($m = 0, 1, \ldots, \eta - 1$) of the n-th symbol interval. It can be expressed as

$$r_{n,m} = \sum_{u=1}^{U} s_{n,m}^{(u)}(\alpha_n^{(u)}, \sigma_n^{(u)}) + w_{n,m} \quad (5)$$

[2]This approach is equivalent of using an orthogonal basis of properly delayed sinc functions.

where, as mentioned, $\{w_{n,m}\}$ are independent and identically distributed complex Gaussian noise samples and $s_{n,m}^{(u)}(\alpha^{(u)}, \sigma^{(u)})$ (whose dependence on $\alpha^{(u)}$ and $\sigma^{(u)}$ will be omitted in the following) is the contribution of user u to the useful signal component. In the following, we will define $\mathbf{r}_n = (r_{n,0}, r_{n,1}, \ldots, r_{n,\eta-1})^T$, $\mathbf{r} = (\mathbf{r}_0^T, \mathbf{r}_1^T, \ldots, \mathbf{r}_{N-1}^T)^T$ and $\mathbf{s}_n^{(u)} = (s_{n,0}^{(u)}, s_{n,1}^{(u)}, \ldots, s_{n,\eta-1}^{(u)})^T$.

III. INFORMATION-THEORETIC ANALYSIS

In this section, we describe a framework for evaluating, from an information-theoretic point of view, the ultimate performance limits of FDM-CPM systems. The analysis is based on the computation of information rate (IR) and spectral efficiency (SE) of FDM-CPM systems and is an extension of the approach described in [4]. The described framework also allows to find the optimal spacing between adjacent channels.

As mentioned, we assume that all users transmit at the same power, employ the same modulation format and are equally spaced in frequency. Under these conditions, the frequency spacing is a measure of the signal bandwidth and the SE can thus be computed. In order to avoid boundary effects, we assume $U \to \infty$. However, for complexity reasons we consider a multiuser detector that assumes the presence of only U' users and treats the other remaining users as additional noise. In other words, the channel model assumed by the receiver is

$$r(t) = \sum_{u=1}^{U'} s^{(u)}(t, \boldsymbol{\alpha}^{(u)}) \exp\{j2\pi f^{(u)}t\} + n(t) \quad (6)$$

where $n(t)$ is a zero-mean circularly-symmetric white Gaussian process with PSD $2(N_0 + N_I)$, N_I being a design parameter optimized through computer simulations [4].

Our goal is to evaluate the ultimate performance limits achievable by a receiver designed for the auxiliary channel (6) when the actual channel is that in (1) with $U \to \infty$. The problem is an instance of mismatched decoding [17] and can be solved by means of the simulation-based method described in [18], which only requires the existence of an algorithm for exact maximum a posteriori (MAP) symbol detection over the auxiliary channel (the optimal multiuser MAP symbol detector for U' users over AWGN [5], [6]).

We first compute the IR for user u as

$$I(\boldsymbol{\alpha}^{(u)}; \mathbf{r}) = \lim_{N \to \infty} \frac{1}{N} E\left\{ \log \frac{p(\mathbf{r}|\boldsymbol{\alpha}^{(u)})}{p(\mathbf{r})} \right\} \quad \left[\frac{\text{b}}{\text{ch. use}}\right]. \quad (7)$$

The probability density functions $p(\mathbf{r}|\boldsymbol{\alpha}^{(u)})$ and $p(\mathbf{r})$ are computed by a forward recursion of the optimal multiuser MAP symbol detector. In (7), the expectation is with respect to the input and output sequences generated according to the model (1). Assuming a system with an infinite number of users, the IR in (7) does not depend on u and we can focus on a generic user. Moreover, we can define the system bandwidth as the separation between two adjacent channels $F = |f^{(u)} - f^{(u-1)}|$ and use it in the definition of the achievable SE

$$\text{SE} = \frac{1}{FT} I(\boldsymbol{\alpha}^{(u)}; \mathbf{r}) \quad \left[\frac{\text{b}}{\text{s} \cdot \text{Hz}}\right]. \quad (8)$$

We remark that the use of the optimal multiuser detector for U' users is necessary to obtain achievable lower bounds on IR and SE. On the other hand, since this optimal multiuser receiver is not interesting for a complexity point of view, in the following we will consider suboptimal reduced-complexity algorithms but still starting from the information-theoretic results on the most efficient modulation formats and the relevant spacings. In addition, we will consider practical coding schemes based on serial concatenation and iterative detection/decoding, and also schemes with $U = U'$ and a limited number of users (at most 5)—spacings, modulation formats, and values of spectral efficiencies obtained from the information-theoretic analysis still represent a very good guideline for the design of practical systems, as shown in the numerical results. Obviously, some (limited) degradation must be expected.

IV. REDUCED-COMPLEXITY MUD ALGORITHM

A. Synchronous Users

By using the FG/SPA framework, we now describe a MUD algorithm without resorting to a Gaussian approximation of the MAI, as done in [12]–[14]. This algorithm has a better performance in this scenario where, in practice, the main contribution of the interference affecting a given user is related to the two adjacent channels only and, hence, the central limit theorem cannot be advocated as in CDMA.

From (5) and assuming that symbols of user u are independent, the pmf $P(\boldsymbol{\alpha}, \boldsymbol{\sigma}|\mathbf{r})$ factorizes as

$$P(\boldsymbol{\alpha}, \boldsymbol{\sigma}|\mathbf{r}) \propto p(\mathbf{r}|\boldsymbol{\alpha}, \boldsymbol{\sigma}) P(\boldsymbol{\sigma}|\boldsymbol{\alpha}) P(\boldsymbol{\alpha})$$

where

$$P(\boldsymbol{\alpha}) = \prod_{n=0}^{N-1} P(\boldsymbol{\alpha}_n) \tag{9}$$

$$P(\boldsymbol{\sigma}|\boldsymbol{\alpha}) = P(\boldsymbol{\sigma}_0) \prod_{n=0}^{N-1} P(\boldsymbol{\sigma}_{n+1}|\boldsymbol{\sigma}_n, \boldsymbol{\alpha}_n) \tag{10}$$

$$p(\mathbf{r}|\boldsymbol{\alpha}, \boldsymbol{\sigma}) \propto \prod_{n=0}^{N-1} \exp\left\{ -\frac{1}{2\Xi^2} \left\| \mathbf{r}_n - \sum_{u=1}^{U} \mathbf{s}_n^{(u)} \right\|^2 \right\}. \tag{11}$$

Notice that $P(\boldsymbol{\sigma}_{n+1}|\boldsymbol{\sigma}_n, \boldsymbol{\alpha}_n)$ is an indicator function, equal to one if $\boldsymbol{\alpha}_n$, $\boldsymbol{\sigma}_n$, and $\boldsymbol{\sigma}_{n+1}$ satisfy the trellis constraints of each user and to zero otherwise. We can further factorize (9), (10), and (11) as

$$P(\boldsymbol{\alpha}) = \prod_{u=1}^{U} \prod_{n=0}^{N-1} P(\alpha_n^{(u)})$$

$$P(\boldsymbol{\sigma}|\boldsymbol{\alpha}) = \prod_{u=1}^{U} P(\sigma_0^{(u)}) \prod_{n=0}^{N-1} P(\sigma_{n+1}^{(u)}|\sigma_n^{(u)}, \alpha_n^{(u)})$$

$$= \prod_{u=1}^{U} P(\sigma_0^{(u)}) \prod_{n=0}^{N-1} I_n^{(u)}(\alpha_n^{(u)}, \sigma_n^{(u)}, \sigma_{n+1}^{(u)})$$

$$p(\mathbf{r}|\boldsymbol{\alpha}, \boldsymbol{\sigma}) \propto \prod_{n=0}^{N-1} F_n(\boldsymbol{\alpha}_n, \boldsymbol{\sigma}_n) \prod_{u=1}^{U} H_n^{(u)}(\alpha_n^{(u)}, \sigma_n^{(u)})$$

where

$$I_n^{(u)}(\alpha_n^{(u)}, \sigma_n^{(u)}, \sigma_{n+1}^{(u)}) = P(\sigma_{n+1}^{(u)}|\sigma_n^{(u)}, \alpha_n^{(u)}) \tag{12}$$

$$H_n^{(u)}(\alpha_n^{(u)}, \sigma_n^{(u)}) = \exp\left\{ \frac{1}{\Xi^2} \mathrm{Re}\left[\mathbf{r}_n^H \mathbf{s}_n^{(u)} \right] \right\} \tag{13}$$

$$F_n(\boldsymbol{\alpha}_n, \boldsymbol{\sigma}_n) = \prod_{i=1}^{U-1} \prod_{j=i+1}^{U} \exp\left\{ -\frac{1}{\Xi^2} \mathrm{Re}\left[\mathbf{s}_n^{(i)H} \mathbf{s}_n^{(j)} \right] \right\}, \tag{14}$$

having discarded the terms independent of symbols and states and taken into account that a CPM signal has a constant envelope. Hence, we finally have

$$P(\boldsymbol{\alpha}, \boldsymbol{\sigma}|\mathbf{r}) \propto \left[\prod_{u=1}^{U} P(\sigma_0^{(u)}) \right] \prod_{n=0}^{N-1} F_n(\boldsymbol{\alpha}_n, \boldsymbol{\sigma}_n)$$

$$\cdot \prod_{u=1}^{U} H_n^{(u)}(\alpha_n^{(u)}, \sigma_n^{(u)}) I_n^{(u)}(\alpha_n^{(u)}, \sigma_n^{(u)}, \sigma_{n+1}^{(u)}) P(\alpha_n^{(u)}). \tag{15}$$

The resulting graph has cycles of length four, that make unlikely the convergence of the SPA, since they are too short. We can remove these short cycles in the original graph by stretching the variables $\sigma_n^{(u)}$ in $(\alpha_n^{(u)}, \sigma_n^{(u)})$. The FG corresponding to this operation has shortest cycles of length twelve. Obviously the SPA applied to this graph is iterative and leads to an approximate computation of the a posteriori probabilities $P(\alpha_n^{(u)}|\mathbf{r})$. We can introduce a further simplification, assuming that the interference among non-adjacent users is negligible. In other words, we approximate (14) as

$$F_n(\boldsymbol{\alpha}_n, \boldsymbol{\sigma}_n) \simeq \prod_{i=1}^{U-1} F_n^{(i,i+1)}(\alpha_n^{(i)}, \sigma_n^{(i)}, \alpha_n^{(i+1)}, \sigma_n^{(i+1)}) \tag{16}$$

where

$$F_n^{(i,i+1)}(\alpha_n^{(i)}, \sigma_n^{(i)}, \alpha_n^{(i+1)}, \sigma_n^{(i+1)})$$

$$= \exp\left\{ -\frac{1}{\Xi^2} \mathrm{Re}\left[\mathbf{s}_n^{(i)H} \mathbf{s}_n^{(i+1)} \right] \right\}. \tag{17}$$

The corresponding FG is shown in Fig. 1.

Let us consider the FG in Fig. 1. If we remove the factor nodes $F_n^{(i,i+1)}$, we obtain U single-user detectors which neglect the interference. Hence, these nodes are in charge of the interference mitigation. The computational aspects of the described algorithm were extensively discussed in [5], [6] where it is shown that it not only has a better performance with respect to the algorithms borrowed from the CDMA literature, but also a much lower complexity.

For the described suboptimal multiuser detector, the performance also depends on the adopted schedule. Serial or parallel schedules are usually considered in the literature [14]. For a lack of space, and also because the difference in performance is practically negligible in the scenario considered in the numerical results of users transmitting at the same power, we only consider the parallel schedule. In this case, at each iteration all users are activated simultaneously. The computed soft-outputs are then provided to the other users for the next iteration and, after deinterleaving, to the decoders. Note that the chosen schedule impacts the latency but not the complexity of the algorithm.
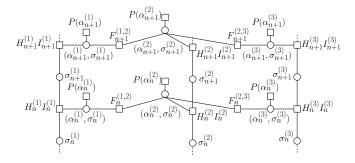
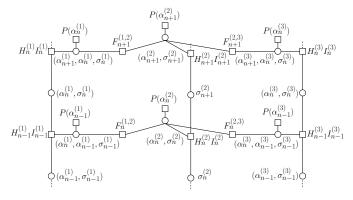Fig. 1. FG resulting from the approximation (16) and for $U = 3$.

Fig. 2. FG for three asynchronous users.

B. Asynchronous Users

The continuous-time model in (1) can be easily modified to account for asynchronous users by including their relative delays. The complex envelope of the received signal is now expressed by

$$r(t) = \sum_{u=1}^{U} s^{(u)}(t - \tau^{(u)}, \boldsymbol{\alpha}^{(u)}) \exp\{j2\pi f^{(u)}(t - \tau^{(u)})\} + w(t)$$

where $\tau^{(u)} \in [0, T)$ is the relative time offset of user u, assumed known at the receiver. Without loss of generality, the smallest of these delays can be assumed to be zero.

The extension of the proposed algorithm to the more general case of asynchronous users is based on the observation that, in this case, the signals of different users are not time-aligned and the interference mitigation stage involves CPM signals on two successive symbol intervals. In other words, while in the synchronous case the interference mitigation can focus on one symbol interval, the nodes $F_n^{(i,j)}$ that connect the pair of users (i, j) in the FG now depend on symbols and states of two consecutive symbol intervals. An example of FG for a system with three asynchronous users with $\tau^{(2)} = 0$ is depicted in Fig. 2, showing how to stretch information symbols. It has been obtained neglecting the interference among non-adjacent channels.

C. Simplified Detection Based on Laurent Decomposition

Laurent decomposition [19] allows to exactly express a CPM signal as the superposition of $(M-1)2^{(L-1)\log_2 M}$ linearly modulated components. Most of the signal power is concentrated in the so-called $M-1$ *principal components*. Hence, reduced-complexity receivers can be designed starting

from an approximation of a CPM signal based on these components only [20], [21]. The proposed MUD algorithm can be redesigned based on an approximate CPM representation using the principal components. The details are omitted for a lack of space. This approach allows to further reduce the computational complexity with a limited performance degradation.

V. CARRIER SYNCHRONIZATION ALGORITHMS

A. Multiuser Joint Detection and Phase Synchronization

We first describe phase synchronization. In the presence of a time-varying phase noise (PN), phase synchronization must be performed jointly with detection [21], [22]. We briefly outline the extension of the proposed reduced-complexity MUD scheme to the case of channels affected by PN. For simplicity, we consider the case of synchronous users. Denoting by $\theta^{(u)}(t)$ the time-varying PN affecting the u-th signal, assumed constant over an interval of length T, the m-th received sample in the n-th symbol interval can be expressed as

$$r_{n,m} = \sum_{u=1}^{U} s_{n,m}^{(u)}(\alpha_n^{(u)}, \sigma_n^{(u)})e^{j\theta_n^{(u)}} + w_{n,m} \qquad (18)$$

where $\theta_n^{(u)} = \theta^{(u)}(nT)$. We follow the Bayesian approach employed in [21] to design single-user detectors for the PN channel. The discrete-time random process $\{\theta_n^{(u)}\}$ is modeled according to a discrete-time Wiener process, whose incremental standard deviation over a symbol interval σ_Δ is known at the receiver [21]. We also assume that $\{\theta_n^{(u)}\}$ and $\{\theta_n^{(v)}\}$, $u \neq v$, are independent of each other.

We can rewrite the signal of user u highlighting the component that depends on the CPM phase state:

$$s_{n,m}^{(u)}(\alpha_n^{(u)}, \sigma_n^{(u)}) = \bar{s}_{n,m}^{(u)}(\alpha_n^{(u)}, \omega_n^{(u)})e^{j\phi_n^{(u)}}.$$

Defining $\psi_n^{(u)} = [\phi_n^{(u)} + \theta_n^{(u)}]_{2\pi}$, the received signal (18) can be expressed as

$$r_{n,m} = \sum_{u=1}^{U} \bar{s}_{n,m}^{(u)}(\alpha_n^{(u)}, \omega_n^{(u)})e^{j\psi_n^{(u)}} + w_{n,m}. \qquad (19)$$

Let us now define $\boldsymbol{\omega}_n = (\omega_n^{(1)}, \ldots, \omega_n^{(U)})^T$, $\boldsymbol{\omega} = (\boldsymbol{\omega}_0^T, \ldots, \boldsymbol{\omega}_N^T)^T$, $\boldsymbol{\psi}_n = (\psi_n^{(1)}, \ldots, \psi_n^{(U)})^T$, $\boldsymbol{\psi} = (\boldsymbol{\psi}_0^T, \ldots, \boldsymbol{\psi}_N^T)^T$, and $\bar{\mathbf{s}}_n^{(u)} = (\bar{s}_{n,0}^{(u)}, \bar{s}_{n,1}^{(u)}, \ldots, \bar{s}_{n,\eta-1}^{(u)})^T$. Discarding irrelevant terms, the joint distribution $p(\boldsymbol{\alpha}, \boldsymbol{\omega}, \boldsymbol{\psi}|\mathbf{r})$ can be factorized as

$$p(\boldsymbol{\alpha}, \boldsymbol{\omega}, \boldsymbol{\psi}|\mathbf{r}) \propto \Big[\prod_{u=1}^{U} P(\omega_0^{(u)})P(\psi_0^{(u)}) \Big] \prod_{n=0}^{N-1} E_n(\boldsymbol{\alpha}_n, \boldsymbol{\omega}_n, \boldsymbol{\psi}_n)$$

$$\cdot \prod_{u=1}^{U} T_n^{(u)}(\alpha_n^{(u)}, \omega_n^{(u)}, \psi_n^{(u)})p(\psi_{n+1}^{(u)}|\psi_n^{(u)}, \alpha_{n-L+1}^{(u)})$$

$$\cdot P(\omega_{n+1}^{(u)}|\omega_n^{(u)}, \alpha_n^{(u)})P(\alpha_n^{(u)}) \quad (20)$$

where

$$T_n^{(u)}(\alpha_n^{(u)}, \sigma_n^{(u)}) = \exp\left\{\frac{1}{\Xi^2}\mathrm{Re}\left[\mathbf{r}_n^H \bar{\mathbf{s}}_n^{(u)} e^{J\psi_n^{(u)}}\right]\right\}$$

$$E_n(\boldsymbol{\alpha}_n, \boldsymbol{\omega}_n, \boldsymbol{\psi}_n) =$$
$$\prod_{i=1}^{U-1}\prod_{k=i+1}^{U} \exp\left\{-\frac{1}{\Xi^2}\mathrm{Re}\left[\bar{\mathbf{s}}_n^{(i)H} \bar{\mathbf{s}}_n^{(k)} e^{-J(\psi_n^{(i)}-\psi_n^{(k)})}\right]\right\}.$$

In (20), $P(\omega_{n+1}^{(u)}|\omega_n^{(u)}, \alpha_n^{(u)})$ is an indicator function, equal to one if $\alpha_n^{(u)}$, $\omega_n^{(u)}$, and $\omega_{n+1}^{(u)}$ are compatible and to zero otherwise, and $p(\psi_{n+1}^{(u)}|\psi_n^{(u)}, \alpha_{n-L+1}^{(u)}) = p(\psi_{n+1}^{(u)}|\psi_n^{(u)}, \omega_n^{(u)})$ is a Gaussian probability density function (pdf) in $\psi_{n+1}^{(u)}$ with mean $[\psi_n^{(u)} + \pi h \alpha_{n-L+1}]_{2\pi}$ and standard deviation σ_Δ. A FG similar to that in Fig. 1 results (under the same assumptions), with proper changes in the labels of variable nodes and functions of factor nodes. A major difference is represented here by the fact that continuous variables $\psi_n^{(u)}$ are now represented in the graph. Hence, the application of the SPA involves computations of continuous pdfs and is not suited for a practical implementation. To overcome this problem, we may resort, as in [21], to the *canonical distribution* approach. Examples of commonly used canonical distributions for this channel can be found in [21]. In the numerical results, we will consider a canonical distribution composed of a weighted sum of impulses. In other words, each phase $\psi_n^{(u)}$ is quantized to D equally spaced values.

Although the algorithm has been obtained by assuming a Wiener PN with known incremental variance over a symbol variance, it can be employed even when the PN follows a different model. In this case, the value of σ_Δ^2 assumed at the receiver must be optimized by simulation for the PN at hand. In any case, there is in general a benefit from using at the receiver a value of thermal noise variance σ^2 larger than the actual one. The rationale of this trick is the following: since there is an overconfidence in the computed messages, we can make the algorithm less confident simply by describing the channel as if it added more noise than it really does [23].

B. Data-Aided Multiuser Fine Frequency Synchronization

The MUD algorithms described in this paper require the knowledge of the amplitudes $\sqrt{2E_S^{(u)}/T}$ and frequency values $f^{(u)}$ for each user. For them, we resort to data-aided (DA) estimation algorithms based on known-data fields usually inserted in the frame. Amplitude estimation is not an issue. In fact, the application of U occurrences of a DA maximum-likelihood single-user estimation algorithm provides amplitude estimates with a good accuracy for typical preamble lengths. Instead, DA single-user frequency estimation algorithms do not provide the required accuracy. This is obviously due to the interference of adjacent channels. For this reason, we employ interference cancellation to refine the estimates.

A first set of estimates of the frequency values $f^{(u)}$ is obtained by applying the DA algorithm in [24] to the preamble of each user. This algorithm does not require the knowledge of the channel phase for each user. These estimates are then iteratively refined still using the same single-user algorithm to the received signal after the contribution of the adjacent signals has been removed. To perform interference cancellation we need to employ not only the already estimated amplitude values and the frequency values of the previous iteration, but also the instantaneous (in case of a time-varying channel phase) values of the channel phase for each user. These are obtained by using the DA multiuser carrier phase estimation algorithm described in the next section, refined every time a new set of frequency estimates becomes available.

In summary, the algorithm proceeds as follows. The amplitude of each user is estimated first. Then, at each iteration a new set of frequency estimates is derived by using the single-user DA algorithm in [24] after the contribution of adjacent users has been removed. This set of frequency estimates is employed to perform DA multiuser carrier phase estimation whose output will be employed for interference cancellation at the next iteration.

A few iterations are in general sufficient, provided the known-data fields of all users have been properly optimized. The results shown in Section VI refer to a properly designed optimization method, not discussed here for a lack of space.

C. Data-Aided Multiuser Carrier Phase Estimation

We now describe a DA multi-user carrier phase estimation algorithm that requires the knowledge of the frequency and amplitude values of each user, estimated as described in the previous section. As mentioned, the phase estimates are used for interference cancellation necessary to improve the frequency estimates.

Let us assume a known-data field of P symbols ($K = \eta P$ samples). Defining $z_k = r_{n,m}$, $x_k^{(u)} = s_{n,m}^{(u)}$, and $\zeta_k = w_{n,m}$, with $k = n\eta + m$, we will assume that the known-data field corresponds to values $k = 0, 1, \ldots, K-1$. We also remove the hypothesis that the PN is constant over a symbol interval and define $\varphi_k^{(u)} = \theta^{(u)}(kT/\eta)$. Hence, we may express

$$z_k = \sum_{u=1}^{U} x_k^{(u)} e^{J\varphi_k^{(u)}} + \zeta_k. \tag{21}$$

Let us define $\boldsymbol{\varphi}_k = (\varphi_k^{(1)}, \ldots, \varphi_k^{(U)})^T$, $\boldsymbol{\varphi} = (\boldsymbol{\varphi}_0^T, \ldots, \boldsymbol{\varphi}_{K-1}^T)^T$ and $\mathbf{z} = (z_0, \ldots, z_{K-1})^T$. As before, we model the PN as a discrete-time Wiener process with incremental standard deviation over a symbol interval σ_Δ. We derive the MAP DA phase estimator:

$$\hat{\varphi}_k^{(u)} = \arg\max_{\varphi_k^{(u)}} p(\varphi_k^{(u)}|\mathbf{z}) \tag{22}$$
$$u = 1, \ldots, U \quad, \quad k = 0, \ldots, K-1.$$

Pdfs $p(\varphi_k^{(u)}|\mathbf{z})$ are obtained from $p(\boldsymbol{\varphi}|\mathbf{z})$ by using the FG/SPA framework. From (21), we may express

$$p(\boldsymbol{\varphi}|\mathbf{z}) \propto p(\mathbf{z}|\boldsymbol{\varphi})p(\boldsymbol{\varphi})$$
$$= \prod_{k=0}^{K-1}\left[p(z_k|\boldsymbol{\varphi}_k)\prod_{u=1}^{U} p(\varphi_k^{(u)}|\varphi_{k-1}^{(u)})\right]$$
$$= \prod_{k=0}^{K-1}\left[p(z_k|\boldsymbol{\varphi}_k)\prod_{u=1}^{U} D_{k,k-1}^{(u)}(\varphi_k^{(u)} - \varphi_{k-1}^{(u)})\right] \tag{23}$$

where $D_{k,k-1}^{(u)}(\varphi_k^{(u)} - \varphi_{k-1}^{(u)}) = p(\varphi_k^{(u)}|\varphi_{k-1}^{(u)})$ is a Gaussian pdf with mean $\varphi_{k-1}^{(u)}$ and standard deviation $\sigma_\Delta/\sqrt{\eta}$, according to
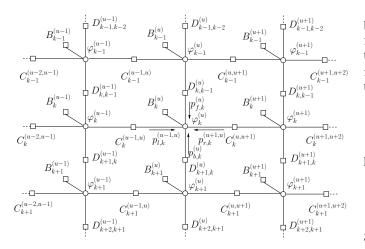
Fig. 3. FG for the multi-user DA phase estimator.

the Wiener model. Neglecting irrelevant multiplicative terms, we can further factorize

$$p(z_k|\boldsymbol{\varphi}_k) \propto \exp\left\{ -\frac{1}{2\sigma^2}\left| z_k - \sum_{u=1}^{U} x_k^{(u)} e^{\jmath\varphi_k^{(u)}} \right|^2 \right\}$$

$$\propto \left(\prod_{u=1}^{U} B_k^{(u)}(\varphi_k^{(u)}) \prod_{u=1}^{U-1}\prod_{v=u+1}^{U} C_k^{(u,v)}(\varphi_k^{(u)},\varphi_k^{(v)}) \right) \tag{24}$$

having defined

$$B_k^{(u)}(\varphi_k^{(u)}) = \exp\left\{ \frac{1}{\sigma^2}\mathrm{Re}\left[z_k x_k^{(u)*} e^{-\jmath\varphi_k^{(u)}} \right] \right\}$$

$$C_k^{(u,v)}(\varphi_k^{(u)},\varphi_k^{(v)}) = \exp\left\{ \frac{1}{\sigma^2}\mathrm{Re}\left[x_k^{(u)} x_k^{(v)*} e^{\jmath(\varphi_k^{(u)}-\varphi_k^{(v)})} \right] \right\}. \tag{25}$$

From (23) and (24), we finally obtain the relevant factorization of $p(\boldsymbol{\varphi}|\mathbf{z})$. Node $C_k^{(u,v)}$ in the resulting factor graph connects variable nodes $\varphi_k^{(u)}$ and $\varphi_k^{(v)}$. Since the interference between non-adjacent users is much smaller than the interference between adjacent users, we consider only functions connecting adjacent variable nodes, i.e., functions $C_k^{(u,u+1)}$. The simplified FG is shown in Fig. 3.

Due to the presence of cycles in the FG of Fig. 3, the application of the SPA gives an iterative algorithm which provides proper approximations of pdfs $p(\varphi_k^{(u)}|\mathbf{z})$. We adopt the *canonical distribution* approach and, as in [22], we model the messages represented in Fig. 3 as Tikhonov pdf, i.e.,

$$p_{f,k}^{(u)}(\varphi_k^{(u)}) = t(a_{f,k}^{(u)};\varphi_k^{(u)})$$
$$p_{b,k}^{(u)}(\varphi_k^{(u)}) = t(a_{b,k}^{(u)};\varphi_k^{(u)})$$
$$p_{l,k}^{(u-1,u)}(\varphi_k^{(u)}) = t(a_{l,k}^{(u-1,u)};\varphi_k^{(u)})$$
$$p_{r,k}^{(u+1,u)}(\varphi_k^{(u)}) = t(a_{r,k}^{(u+1,u)};\varphi_k^{(u)}) \tag{26}$$

where $t(\xi;x)$ is a Tikhonov distribution in x characterized by the complex parameter ξ:

$$t(\xi;x) = \frac{1}{2\pi\mathrm{I}_0(|\xi|)} \exp\{\mathrm{Re}[\xi e^{-jx}]\}$$

being $\mathrm{I}_0(x)$ the zeroth-order modified Bessel function of the first kind. Hence, we simply have to update and propagate the complex parameters describing the Tikhonov pdfs. Let us first consider the update of parameter $a_{f,k}^{(u)}$. By generalizing the results in [22], we have

$$a_{f,k+1}^{(u)} = \gamma\left(a_{f,k}^{(u)} + \frac{z_k x_k^{(u)*}}{\sigma^2} + a_{l,k}^{(u-1,u)} + a_{r,k}^{(u+1,u)}, \sigma_\Delta \right) \tag{27}$$

having defined

$$\gamma(\xi,\sigma_\Delta) = \frac{\xi}{1+|\xi|\sigma_\Delta^2}.$$

Similarly,

$$a_{b,k-1}^{(u)} = \gamma\left(a_{b,k}^{(u)} + \frac{z_k x_k^{(u)*}}{\sigma^2} + a_{l,k}^{(u-1,u)} + a_{r,k}^{(u+1,u)}, \sigma_\Delta \right). \tag{28}$$

Regarding parameter $a_{l,k}^{(u-1,u)}$ and $a_{r,k}^{(u+1,u)}$ we have

$$a_{l,k}^{(u,u+1)} = \delta\left(a_{f,k}^{(u)} + a_{b,k}^{(u)} + \frac{z_k x_k^{(u)*}}{\sigma^2} + a_{l,k}^{(u-1,u)}, \frac{x_k^{(u)} x_k^{(u+1)*}}{\sigma^2} \right) \tag{29}$$

and

$$a_{r,k}^{(u,u-1)} = \delta\left(a_{f,k}^{(u)} + a_{b,k}^{(u)} + \frac{z_k x_k^{(u)*}}{\sigma^2} + a_{r,k}^{(u+1,u)}, \frac{x_k^{(u-1)} x_k^{(u)*}}{\sigma^2} \right) \tag{30}$$

having defined

$$\delta(\xi,\eta) = \frac{\xi\eta}{\sqrt{|\xi|^2+|\eta|^2}}.$$

In order to obtain (29) and (30), two approximations have been employed: $\mathrm{I}_0(|x|) \simeq e^{|x|}$ and $\sqrt{1+x} \simeq 1+x/2$. The following schedule is adopted: messages $a_{f,k}^{(u)}$ and $a_{b,k}^{(u)}$ are first updated, for $k = 0, \ldots, K-1$ (with initial parameter $a_{f,0}^{(u)} = 0$) and $k = K-1, \ldots, 0$ (with initial parameter $a_{b,K-1}^{(u)} = 0$), respectively. Then messages $a_{l,k}^{(u-1,u)}$ and $a_{r,k}^{(u+1,u)}$ are updated for $u = 2, \ldots, U$ and $u = U-1, \ldots, 1$, respectively. A couple of iterations are, in general, sufficient. Finally, the estimates are

$$\hat{\varphi}_k^{(u)} = \arg\left(a_{f,k}^{(u)} + a_{b,k}^{(u)} + a_{l,k}^{(u-1,u)} + a_{r,k}^{(u+1,u)} + \frac{z_k x_k^{(u)*}}{\sigma^2} \right). \tag{31}$$

VI. NUMERICAL RESULTS

For a lack of space, we limit our investigation to binary CPM formats. This choice is justified by the need to illustrate the relevant concepts and by the results which show that we can design transmission schemes with a very high efficiency using simple CPMs. We consider binary CPMs with $h = 1/3$, $L = 2$, and rectangular (REC) or raised-cosine frequency pulse (RC). These schemes turned out to be the best ones among those considered.
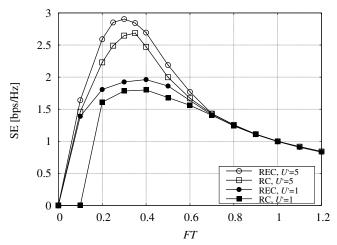
111

Fig. 4. Spectral efficiency as a function of the normalized spacing for different values of U' at $E_s/N_0 = 10$ dB.
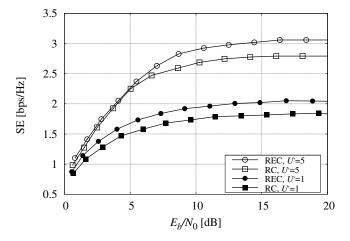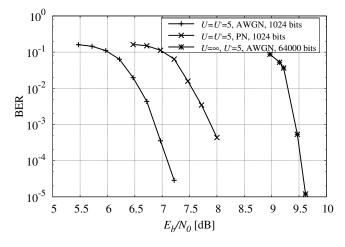

Fig. 5. Spectral efficiency as a function of E_b/N_0, for channel spacings resulting from the optimization at $E_s/N_0 = 10$ dB.

A. Information-Theoretic Analysis

Fig. 4 shows the SE as a function of the normalized spacing FT for the considered CPM formats and $E_s/N_0 = 10$ dB. In particular, we computed the achievable SE when a multiuser detector for different values of U' is employed. Note that curves with $U' = 1$ in practice refer to the use of a single-user detector. For each considered CPM scheme, a multiuser detection strategy allows to achieve a higher SE than that achievable by a single-user detector and the frequency spacing which maximizes the SE is also lower in the multiuser detection case.

As discussed in [4], the optimal spacing depends on the considered value of E_s/N_0 (although this dependence is quite smooth). Hence, according to the operating E_s/N_0, we choose the optimal modulation format and the corresponding optimal spacing. For the REC scheme, $FT = 0.3$ (resp. 0.4) is the optimal spacing at $E_s/N_0 = 10$ dB when $U' = 5$ (resp. $U' = 1$), whereas $FT = 0.35$ (resp. 0.4) is that for RC when $U' = 5$ (resp. $U' = 1$). From Fig. 4, we can also observe that the scheme with REC frequency pulse leads to a higher SE than the RC-based counterpart. The same conclusion can be drawn from Fig. 5, where we consider the channel spacings resulting from the previous optimization and we plot the SE as

Fig. 6. BER performance.

a function of E_b/N_0, being E_b the energy per information bit. The figure shows that REC and RC formats perform similarly for low values of SE. However, REC allows achieving a higher SE.

B. BER Performance

In the following, we will consider the design of a practical scheme based on the REC frequency pulse with a SE of 2.66 bps/Hz, slightly lower than the maximum achievable since we will consider suboptimal reduced-complexity detection schemes. From Fig. 5 a corresponding value of $E_b/N_0 \simeq 7.6$ dB results. To obtain this SE, we concatenate, through an interleaver, the REC scheme with spacing $FT = 0.3$ with a (64,51) outer extended Bose, Ray-Chaudhuri, Hocquenghem (eBCH) code with rate $\rho = 0.8$ described in [25].
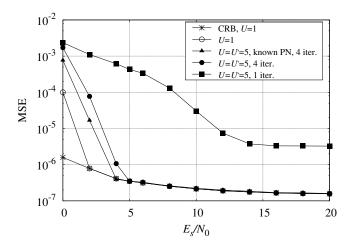
We consider iterative receivers employing the described reduced-complexity multiuser detectors as inner SISO detectors with a maximum of 20 iterations in all cases. The bit error rate (BER) performance for the middle user only is shown in Fig. 6 versus E_b/N_0. We assume that the users are synchronous. However, we also carried out BER analysis in the asynchronous scenario observing the same performance as for the synchronous case. We do not report the performance of the optimal MUD scheme—see [5], [6] for a performance comparison.

When $U = \infty$ and $U' = 5$, thus in the same conditions used in the information-theoretic analysis (although with a different receiver) we considered codewords of length 64000 information bits and the AWGN channel. A convergence threshold of about $E_b/N_0 \simeq 9.2$ dB and BER of 10^{-5} for $E_b/N_0 \simeq 9.6$ dB are obtained, less than a couple of dB worse than predicted. We also show the performance for $U = U' = 5$ and AWGN and PN channels—in the latter case a Wiener model with $\sigma_\Delta = 1$ degree has been considered and at the receiver $D = 24$ has been adopted.

It is worth noticing that the proposed binary schemes outperform the quaternary schemes under consideration for next generation Digital Video Broadcasting - Return Channel Satellite (DVB-RCS) standard.

C. Synchronization

For the above mentioned REC-based scheme we now consider the issue of frequency synchronization. The described

Fig. 7. MSE of the multiuser frequency synchronization scheme in the presence of PN.

multiuser frequency synchronization scheme results unbiased. Hence, in Fig. 7 we show the mean square error (MSE) of the frequency estimate for the central user versus E_S/N_0, when $P = 60$ bits. A Wiener phase noise with $\sigma_\Delta = 1$ degree has been considered. As a reference curve, we show the Cramér-Rao lower bound (CRB) for a system with $U = 1$, computed according to [26]. When $U = 1$, this bound is reached by the frequency estimation algorithm in [24] for $E_S/N_0 \geq 2$ dB (curve with white circles). When $U = U' = 5$ users are present, this algorithm gives a very poor performance (curve labeled "1 iteration" since it corresponds to the first iteration of the proposed multiuser algorithm). With 4 iterations we are able to reach, for $E_S/N_0 \geq 5$ dB, the CRB related to the presence of only one user. Hence, a very effective interference cancellation is performed. A slightly better result is obtained by using a genie-aided version of the proposed frequency synchronization algorithm in which the true values of the channel phases are employed for interference cancellation purposes. This is a proof of the effectiveness of the proposed DA multiuser carrier phase estimation algorithm described in Section V-C.

VII. CONCLUSIONS

We considered frequency division multiplexed systems based on continuous phase modulations. Through the information-theoretic analysis, we showed that it turns out more convenient to consider frequency spacings between the channels much lower than those usually employed to limit the interference from adjacent users, thus providing a better tradeoff between degradation of the information rate due to the interference and usage of the available spectrum. The limits predicted by the information theory can be approached by practical schemes based on the serial concatenation with a proper outer code.

We then proposed reduced-complexity schemes for multiuser detection, possibly in the presence of phase noise, and multiuser data-aided phase and frequency synchronization schemes. We showed that it is possible to implement transmission schemes with an unprecedented spectral efficiency at a price of a limited complexity increase with respect to a receiver which neglects the interference.

REFERENCES

[1] B. F. Beidas, H. El Gamal, and S. Kay, "Iterative interference cancellation for high spectral efficiency satellite communications," *IEEE Trans. Commun.*, vol. 50, pp. 31–36, Jan. 2002.
[2] G. Gallinaro and R. Rinaldo, "Assessment of potentiality of adjacent channel interference mitigation in a low-rate TDMA system," in *Proc. Intern. Work. on Signal Processing for Space Commun.*, (Rhodes Island, Greece), October 2008.
[3] A. Barbieri, D. Fertonani, and G. Colavolpe, "Time-frequency packing for linear modulations: spectral efficiency and practical detection schemes," *IEEE Trans. Commun.*, vol. 57, pp. 2951–2959, Oct. 2009.
[4] A. Barbieri, D. Fertonani, and G. Colavolpe, "Spectrally-efficient continuous phase modulations," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 1564–1572, Mar. 2009.
[5] A. Piemontese and G. Colavolpe, "A novel graph-based soft interference cancellation algorithm for FDM-CPM transmissions," submitted to *IEEE Trans. on Wireless Commun.*, 2009.
[6] A. Piemontese and G. Colavolpe, "A novel graph-based soft interference cancellation algorithm for FDM-CPM satellite systems," in *Proc. IEEE International Symposium on Information Theory*, (Seoul, Korea), June 2009.
[7] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inform. Theory*, vol. 47, pp. 498–519, Feb. 2001.
[8] G. Ferrari, G. Colavolpe, and R. Raheli, *Detection Algorithms for Wireless Communications*. John Wiley & Sons, 2004.
[9] G. Colavolpe and A. Barbieri, "On MAP symbol detection for ISI channels using the Ungerboeck observation model," *IEEE Commun. Letters*, vol. 9, pp. 720–722, Aug. 2005.
[10] K. R. Narayanan and G. L. Stüber, "Performance of trellis-coded CPM with iterative demodulation and decoding," *IEEE Trans. Commun.*, vol. 49, pp. 676–687, Apr. 2001.
[11] P. Moqvist and T. M. Aulin, "Serially concatenated continuous phase modulation with iterative decoding," *IEEE Trans. Commun.*, vol. 49, pp. 1901–1915, Nov. 2001.
[12] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, pp. 1046–1061, July 1999.
[13] H. El Gamal and E. Geraniotis, "Iterative multiuser detection for coded CDMA signals in AWGN and fading channels," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 30–41, Jan. 2000.
[14] J. Boutros and G. Caire, "Iterative multiuser joint decoding: unified framework and asymptotic analysis," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1772–1793, July 2002.
[15] B. E. Rimoldi, "A decomposition approach to CPM," *IEEE Trans. Inform. Theory*, vol. 34, pp. 260–270, Mar. 1988.
[16] H. Meyr, M. Oerder, and A. Polydoros, "On sampling rate, analog prefiltering, and sufficient statistics for digital receivers," *IEEE Trans. Commun.*, vol. 42, pp. 3208–3214, Dec. 1994.
[17] N. Merhav, G. Kaplan, A. Lapidoth, and S. S. Shitz, "On information rates for mismatched decoders," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1953–1967, Nov. 1994.
[18] D. M. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavčić, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Trans. Inform. Theory*, vol. 52, pp. 3498–3508, Aug. 2006.
[19] U. Mengali and M. Morelli, "Decomposition of M-ary CPM signals into PAM waveforms," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1265–1275, Sept. 1995.
[20] G. Colavolpe and R. Raheli, "Noncoherent sequence detection of continuous phase modulations," *IEEE Trans. Commun.*, vol. 47, pp. 1303–1307, Sept. 1999.
[21] A. Barbieri and G. Colavolpe, "Simplified soft-output detection of CPM signals over coherent and phase noise channels," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 2486–2496, July 2007.
[22] G. Colavolpe, A. Barbieri, and G. Caire, "Algorithms for iterative decoding in the presence of strong phase noise," *IEEE J. Select. Areas Commun.*, vol. 23, pp. 1748–1757, Sept. 2005.
[23] G. Colavolpe, G. Ferrari, and R. Raheli, "Extrinsic information in iterative decoding: a unified view," *IEEE Trans. Commun.*, vol. 49, pp. 2088–2094, Dec. 2001.
[24] U. Mengali and M. Morelli, "Data-aided frequency estimation for burst digital transmission," *IEEE Trans. Commun.*, vol. 45, pp. 23–25, Jan. 1997.
[25] A. Graell i Amat, C. A. Nour, and C. Douillard, "Serially concatenated continuous phase modulation for satellite communications," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 3260–3269, June 2009.
[26] A. Barbieri and G. Colavolpe, "On the Cramer-Rao bound for carrier frequency estimation in the presence of phase noise," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 575–582, Feb. 2007.