Generalized Differential Encoding: A Nonlinear Signal Processing Perspective

Fulvio Gini and Georgios B. Giannakis, *Fellow, IEEE*

***Abstract*—Incoherent detection based on differential encoding has been successfully applied to phase-shift-keying (PSK) signals because it eliminates the need for carrier phase acquisition and tracking at the receiver. This paper generalizes the idea of differential encoding by using a nonlinear transformation called multilag high-order instantaneous moment (ml-HIM). The ml-HIM decoder is capable of removing the effects of phase ambiguity, Doppler frequency, Doppler rate, and even higher order phase distortions. The degrees of freedom offered by the different lags is exploited to optimize system performance. In addition to the classical M-ary PSK, the generalized differential encoding idea is also applied to nonconstant modulus constellations, such as M-ary QAM and AM-PM.**

I. Introduction

IN DIGITAL communication systems, phase ambiguity and/or frequency errors are likely to be present in the received signal due to imperfect knowledge of the carrier's phase and frequency, fading effects, or multipath propagation [3, ch. 6]. Typically, phase distortions are captured in a term $\exp[j\theta_c(t)]$, which multiplies the received baseband data [see (1)]. When phase variations are induced by the relative motion between transmitter and receiver (such as in mobile and satellite communications), the phase $\theta_c(t)$ is polynomial in t, and its coefficients are related to the kinematics of the moving station [17, p. 59]. Moreover, thanks to the Stone–Weierstrass theorem, any continuous function over a closed interval can be approximated uniformly by a polynomial function (see e.g., [16, p. 399]). The polynomial coefficients, which are generally modeled as deterministic unknown quantities, are estimated first, and the received data are multiplied next by $\exp[-j\hat{\theta}_c(t)]$ to remove phase distortions (see e.g., [5] for constant phase estimation and [6] and [9] for frequency estimation methods). In optical communications, phase fluctuations are due to instabilities of the transmit laser. In heterodyne reception, they arise due to both transmit and receive local oscillator lasers that have linewidths of tens of megahertz and exhibit instability greater than microwave oscillators [4], [20].

As an alternative to modeling and estimating phase distortions, phase errors can be precompensated by differential encoding at the transmitter and differential decoding at the receiver. The latter eliminates the need for carrier phase acquisition and tracking [3, ch. 5] at the expense of additional signal power required to attain a given probability of error (when compared with the ideal coherent detection). Traditionally, information transmitted with *differential phase-shift-keying* (DPSK) is encoded in phase differences between two successive symbols. Although tolerant to constant phase errors, differential detection systems are sensitive to carrier frequency variations [4], [18].

To overcome this problem, *doubly differential* PSK (DDPSK) has been discussed recently for M-ary PSK in [18], based on [15] and less-known results from the Russian literature [12], [21]. However, in applications such as low earth orbiting (LEO) satellite communications [10], the Doppler frequency also changes with time, and DDPSK is not adequate for compensation. Algorithms for frequency and frequency rate-of-change estimation are proposed in [8], based on Lanczos FIR differentiators, but they suffer from non-negligible estimation bias. Generalizations to higher- than-second-order M-ary DPSK have been also suggested in [19], but no performance results are reported. Another interesting approach has been proposed in [10], where a dual-channel PSK demodulator for LEO satellite direct sequence/code division multiple access (DS/CDMA) communications is proposed, whose performance are tolerant to time-variant Doppler frequency.

Interestingly, differential encoding has been restricted only to MPSK modulation systems, i.e., to constant modulus constellations. In [5], however, blind phase recovery in quadrature amplitude modulated (QAM) signals was proposed based on higher order statistics (HOS) of the received data. Unfortunately, unambiguous estimates are guaranteed only if phase errors lie in $[-\pi/4, \pi/4]$, and processing delay is incurred in order to collect the samples required for accurate HOS estimation [5].

In this paper, we revisit the differential encoding idea from a nonlinear signal processing perspective that relies on the multilag high-order instantaneous moment (ml-HIM), which is a transformation originally introduced in [1] and thoroughly studied in [2]. Starting from the well-known case of MDPSK modulation [3, ch. 5], we show how differential encoding can be interpreted in terms of the ml-HIM (Section II) and how generalized differential encoding techniques can be derived by means of the multi-lag HIM (Section III). In particular, the degrees of freedom offered by the different lags are exploited to increase system performance without increasing complexity. The maximum deflection criterion is used as a tool for selecting the lags (Section IV). This approach

Manuscript received February 28, 1997; revised February 19, 1998. This wotk was supported by ONR under Grant N00014-93-1-0485. The associate editor coordinating the review of this paper and approving it for publication was Dr. Jonathon A. Chambers.

F. Gini is with the Department Ingegneria dell'Informazione, University of Pisa, Pisa, Italy (e-mail: gini@iet.unipi.it).

G. B. Gianakis is with the Department of Electrical Engineering, University of Virginia, Charlottesville, VA 22903-2442 USA (e-mail: georgios@virginia.edu).

Publisher Item Identifier S 1053-587X(98)07806-4.

is particularly useful both when the additive noise is not white, or when higher-than-second-order differential encoding is employed in order to eliminate the Doppler effects present in the received data. Finally, it is shown, in Section V, how differential encoding can be successfully applied to generic (nonconstant modulus) constellations, such as M-ary QAM and *amplitude-modulation phase-modulation* (AM-PM) constellations [3, ch. 5], with the following advantages:

1) no increase in complexity of the modulator structure;
2) no carrier recovery circuit;
3) performance invariance to phase distortions, in particular, invariance to Doppler frequency and Doppler rate.

II. Preliminaries

We consider here transmission of a linearly modulated signal through a known linear channel and assume perfect symbol synchronization at the receiver.

A. Signal Model

The complex envelope of the received signal after baseband conversion can be written as, [3, ch. 5], [6]

$$r_c(t) = e^{j\theta_c(t)} \sum_l w_d(l) g_c^{(\mathrm{tr})}(t - lT_s) + n_c(t) \qquad (1)$$

where

- $\theta_c(t)$ models the composite phase error introduced by the channel and imperfect knowledge of the carrier signal's phase and frequency;
- $g_c^{(\mathrm{tr})}(t)$ cascade of the transmitter's signaling pulse and the channel impulse response;
- T_s symbol period;
- $w_d(l)$s transmitted complex symbols.

If we denote by $w(l)$ the information sequence, we have $w_d(l) = w(n)$ when no encoding is introduced at the transmitter; otherwise, the relationship between information and transmitted symbols depends on the particular encoding algorithm. The noise $n_c(t)$ is assumed complex white Gaussian, with two-sided power spectral $N_0/2$. Subscript c denotes continuous-time signals.

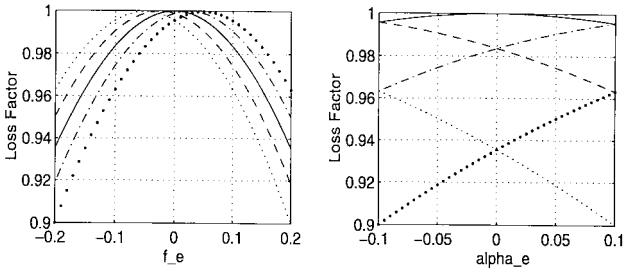
After the receiving matched filter $g_c^{(\mathrm{rec})}(t)$, the signal $x_c(t) = r_c * g_c^{(\mathrm{rec})}(t)$ ($*$ denotes convolution) is sampled at the symbol rate $1/T_s$. We thus obtain the discrete-time data

$$x(n) = e^{j\theta(n)} \sum_l w_d(l) g(n-l) + v(n) \qquad (2)$$

where $x(n) := x_c(nT_s)$, $\theta(n) := \theta_c(nT_s)$, $v(n) := n_c * g_c^{(\mathrm{rec})}(t)|_{t=nT_s}$, and $g(n) := g_c^{(\mathrm{tr})} * g_c^{(\mathrm{rec})}(t)|_{t=nT_s}$ denotes the cascade of the transmit filter, the channel impulse response, and the receive filter. Model (2) is valid as far as the mismatch of the receive filter $g_c^{(\mathrm{rec})}(t)$ due to $\theta_c(t)$ can be neglected, i.e., the bandwidth of $\exp[j\theta_c(t)]$ is small as compared with that of $g_c^{(\mathrm{rec})}(t)$.

Suppose that a second-order approximation represents an accurate model of $\theta_c(t)$

$$\theta_c(t) := \theta_0 + 2\pi f_d t + \pi \alpha t^2 \qquad (3)$$



(a) (b)

Fig. 1. (a) Loss factor $\lambda(f_e, \alpha_e)$ versus f_e for $\alpha_e = -0.1$ (dotted line), -0.05 (dashed), $0.$ (solid), 0.05 (dash-dotted), and 0.1 (tick dotted). (b) $\lambda(f_e, \alpha_e)$ versus α_e for $f_e = -0.1$ (dotted line), -0.05 (dashed), $0.$ (solid), 0.05 (dash-dotted), and 0.1 (tick dotted).

with θ_0 denoting the phase offset, whereas f_d and α are the Doppler shift and Doppler rate, respectively. This assumption implies that the instantaneous Doppler frequency is given by $\dot{\theta}_c(t) = 2\pi(f_d + \alpha t)$, that is, linear in t (see e.g., [10, Fig. 1]) for a typical plot of $\dot{\theta}_c(t)$]. Higher order phase polynomials can be considered as well. With $\theta_c(t)$ given by (3), model (2) holds, provided that Doppler frequency and Doppler rate are small compared with the symbol rate, which is a condition satisfied in many practical cases (see, e.g., [6]). An exception is the LEO satellite system, which suffers from large Doppler shifts ranging from -60 kHz to 60 kHz for a carrier frequency of 2.4 GHz and a symbol rate of 9.6 kb/s. If a filter matched to the signaling pulse is used, due to the Doppler shift, the received signal would occupy a frequency range disjoined from that of the receive filter, with a consequent loss of signal energy (that is filtered out). A filter bandwidth of approximately six times wider than the symbol rate is needed for (2) to be valid by avoiding energy loss and intersymbol interference (ISI) [10]. To further investigate the effect of f_d and α, consider a rectangular pulse of duration T_s so that $g_c^{(\mathrm{rec})}(t) = g_c^{(\mathrm{tr})}(T_s - t) = T_s^{-1/2}\mathrm{rect}(t/T_s)$. The exact expression of $x(n)$, in the noise-free case, is given by

$$\begin{aligned} x(n) &= e^{j(\theta_0 + 2\pi f_d n T_s + \pi\alpha n^2 T_s^2)} \sum_l w_d(l) \\ &\quad \times \frac{1}{T_s} \int_{-\infty}^{+\infty} \mathrm{rect}\left(\frac{\tau - lT_s}{T_s}\right) \\ &\quad \times \mathrm{rect}\left(\frac{\tau - nT_s}{T_s}\right) e^{-j2\pi[(f_d + n\alpha T_s)\tau - \alpha T_s \tau^2/2]}\, d\tau \\ &= e^{j(\theta_0 + 2\pi f_e n + \pi\alpha_e n^2)} w_d(n) \frac{1}{T_s} \\ &\quad \times \int_0^{T_s} e^{-j2\pi[(f_d + n\alpha T_s)\tau - \alpha T_s \tau^2/2]}\, d\tau \\ &= e^{j(\theta_0 + 2\pi f_e n + \pi\alpha_e n^2)} w_d(n) \lambda(f_e, \alpha_e) \end{aligned}$$

where we defined $\lambda(f_e, \alpha_e) := \int_0^1 \exp\{-j2\pi[(f_e + n\alpha_e)y - \alpha_e y^2/2]\}\, dy$, $y := \tau/T_s$, $f_e := f_d T_s$, and $\alpha_e := \alpha T_s^2$. In Fig. 1(a), we plot the loss factor $|\lambda(f_e, \alpha_e)|$ versus f_e for different values of α_e, whereas in Fig. 1(b), $|\lambda(f_e, \alpha_e)|$ is plotted as a function of α_e for various f_e's. The results show that for $f_e < 0.1$ and $\alpha_e < 0.05$, we can consider $\lambda(f_e, \alpha_e) \approx 1$, and this is the case in most applications (see also [6] and [18] for further details).

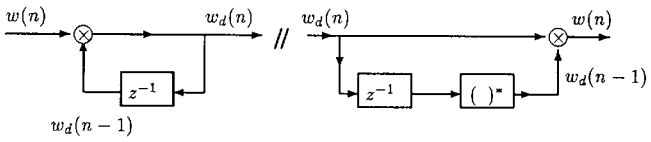


Fig. 2. First-order differential encoder/decoder $(m_1 = 1)$.

In conclusion, sampling at $t_n = nT_s$ to satisfy the Nyquist condition, we have $g(n) = \delta(n)$ [3, p. 326], where $\delta(n)$ is the Kronecker delta function. The received sequence is then

$$x(n) = w_d(n)e^{j(\theta_0 + 2\pi f_e n + \pi\alpha_e n^2)} + v(n). \tag{4}$$

Noise $v(n)$ is zero-mean complex Gaussian with autocorrelation sequence

$$\begin{aligned} R_v(l) &:= E\{v(n)v^*(n+l)\} \\ &= \frac{N_0}{2} g_c^{(\mathrm{rec})}(t) * g_c^{(\mathrm{rec})}(-t)|_{t=lT_s} \\ &= \frac{N_0}{2} g(l) = \frac{N_0}{2}\delta(l) \end{aligned}$$

so that $v(n)$ is white with variance $\sigma_v^2 := N_0/2$. Equation (4) represents our discrete-time received signal model.

B. The Multilag HIM

At this point, it is useful to recall the definition of ml-HIM. The ml-HIM of $x(n)$ is defined recursively as[1]

$$\begin{aligned} x_1(n) &:= x(n) \\ x_2(n; m_1) &:= x_1(n)x_1^*(n - m_1)\ldots \\ x_k(n; m_1, m_2, \ldots, m_{k-1}) &:= x_{k-1}(n; m_1, \ldots, m_{k-2})x_{k-1}^* \\ &\quad \times (n - m_{k-1}; m_1, \ldots, m_{k-2}). \end{aligned} \tag{5}$$

We term x_k the kth order multilag HIM because it reduces to the HIM for $m_1 = m_2 = \cdots = m_{k-1}$, [16, ch. 12], [14], [22]. To show the role played by the ml-HIM in differential encoding, let us consider first $\theta(n) = \theta_0$ and M-ary DPSK symbols.

C. An Example: MDPSK Signaling

The information sequence $w(n) := \exp[j\psi(n)]$ is assumed to be i.i.d., drawn from a discrete M-ary alphabet set $\{\exp[j(2\pi k/M + \psi)]\}_{k=0}^{M-1}$ with probability $1/M$ each; without loss of generality, we assume $\psi = 0$. Let $w_d(n) := \exp[j\psi_d(n)]$ be the differentially encoded sequence. Because the information is encoded in phase differences between two successive samples, we have $\psi(n) = \psi_d(n) - \psi_d(n-1)$, which implies [see also Fig. 2(a)]

$$w_d(n) = e^{j\psi_d(n)} = e^{j[\psi(n) + \psi_d(n-1)]} = w(n)w_d(n-1). \tag{6}$$

Equation (6) represents the input–output relationship of the encoder at the transmitter site.

[1] This definition is slightly different from the one in [2] in order to guarantee the causality of the transformation so that it can be implemented *on line* by the decoder.

From (4), the received discrete-time signal is given by

$$x(n) = w_d(n)e^{j(\theta_0 + 2\pi f_e n + \pi\alpha_e n^2)} + v(n). \tag{7}$$

At the receiver, the information sequence $w(n)$ can be recovered by "inverting" (6)

$$w(n) = e^{j\psi(n)} = e^{j[\psi_d(n) - \psi_d(n-1)]} = w_d(n)w_d^*(n-1). \tag{8}$$

In the noise-free case, if $f_e = 0$ and $\alpha_e = 0$, $w(n)$ is decoded using (8). When noise is present, the decoder output must be quantized, e.g., in the simple binary case (BDPSK) $w(n) \in \{-1, 1\}$, the transmitted sequence is estimated via

$$\hat{w}(n) = \mathrm{sgn}\{x(n)x^*(n-1)\} \tag{9}$$

where $\mathrm{sgn}\{\cdot\}$ denotes the signum function of its argument. The encoding-decoding strategy in (7)–(9) is not affected by a phase shift θ_0, and in the noise-free case, we have

$$\hat{w}(n) = x(n)x^*(n-1) = w_d(n)w_d^*(n-1) = w(n). \tag{10}$$

Equation (8) shows that the decoder implements nothing but the second-order instantaneous moment defined in (6) and evaluated for $m_1 = 1$. Thus, (9) tells us that the decision strategy consists of properly quantizing the second-order HIM of the observed signal.

This observation suggests how to generalize the idea of differential encoding to the case where $\theta(n)$ is not simply a constant θ_0. It also allows us to exploit useful properties of the multilag HIM transformation, namely, its unbiasedness and consistency [22]. The goal is to encode $w(n)$ to $w_d(n)$ so that $\theta(n)$ in (4) does not affect recovery of $w_d(n)$ from $x(n)$ samples.

III. Differential Decoding via Multilag HIM

We show next how the multilag HIM can be applied to remove Doppler frequency and Doppler rate effects (and higher phase distortions if necessary). Consider the ml-HIM (up to the fourth order) of the noise-free signal in (7):

$$x_1(n) = x(n) = w_d(n)e^{j(\theta_0 + 2\pi f_e n + \pi\alpha_e n^2)} \tag{11}$$

$$\begin{aligned} x_2(n; m_1) &= x_1(n)x_1^*(n - m_1) \\ &= w_d(n)w_d^*(n - m_1) \\ &\quad \times e^{j(2\pi f_e m_1 - \pi\alpha_e m_1^2 + 2\pi\alpha_e m_1 n)} \end{aligned} \tag{12}$$

$$\begin{aligned} x_3(n; m_1, m_2) &= x(n)x^*(n - m_1)x^*(n - m_2) \\ &\quad \times x(n - m_1 - m_2) \\ &= w_d(n)w_d^*(n - m_1)w_d^*(n - m_2) \\ &\quad \times w_d(n - m_1 - m_2)e^{j2\pi\alpha_e m_1 m_2} \end{aligned} \tag{13}$$

$$\begin{aligned} x_4(n; m_1, m_2, m_3) &= x(n)x^*(n - m_1)x^*(n - m_2) \\ &\quad \times x^*(n - m_3)x(n - m_1 - m_2) \\ &\quad \times x(n - m_1 - m_3)x(n - m_2 - m_3) \\ &\quad \times x^*(n - m_1 - m_2 - m_3) \\ &= w_d(n)w_d^*(n - m_1) \\ &\quad \times w_d^*(n - m_2)w_d^*(n - m_3) \\ &\quad \times w_d(n - m_1 - m_2) \\ &\quad \times w_d(n - m_1 - m_3)w_d(n - m_2 - m_3) \\ &\quad \times w_d^*(n - m_1 - m_2 - m_3). \end{aligned} \tag{14}$$

These relationships describe the algorithm implemented by the decoder. For example, (12) shows that to remove a constant phase ambiguity, it is sufficient for the decoder to implement the second-order ml-HIM of the received data, but this does not remove Doppler frequency and/or Doppler rate. Note that (8) corresponds to (12) evaluated for $m_1 = 1$. The third-order ml-HIM can remove both constant phase and Doppler frequency but not Doppler rate, for which it is necessary to resort to the fourth-order ml-HIM. Higher order terms in the power expansion of $\theta(n)$ require implementation of higher order ml-HIM to be eliminated.

Generalizing (6), the encoder at the transmitter should correspondingly implement one the following input–output relationships:

$$w_d(n) = w(n)w_d(n-m_1) \tag{15}$$

$$w_d(n) = w(n)w_d(n-m_1)w_d(n-m_2)w_d^*(n-m_1-m_2) \tag{16}$$

$$\begin{aligned} w_d(n) = {} & w(n)w_d(n-m_1)w_d(n-m_2)w_d(n-m_3) \\ & \times w_d^*(n-m_1-m_2)w_d^*(n-m_1-m_3) \\ & \times w_d^*(n-m_2-m_3)w_d(n-m_1-m_2-m_3). \end{aligned} \tag{17}$$

To assure causality, we select $0 < m_1 \leq m_2 \leq m_3$. Inductively, to remove a Qth-order polynomial phase distortion $\theta(n)$, we must encode with the kth-order ml-HIM and select $k = Q + 2$. Under the white noise assumption, $x_k(n; m_1, \ldots, m_{k-1})$ is known to be an unbiased and mean-square sense consistent estimator of the kth order ml-HIM of $w_d(n)$ [22], which coincides with $w(n)$ if the transmitted symbols are properly encoded according to (15)–(17).

Taking into account that $w(n)$ comes from a finite alphabet, the receiver will choose the symbol $w(n)$ that is closest to its noisy estimate $x_k(n; m_1, \ldots, m_{k-1})$; thus, the decision rule is given by

$$\hat{w}(n) = \arg\min_{w(n)} |w(n) - x_k(n; m_1, \ldots, m_{k-1})|. \tag{18}$$

The decoding strategy of (13) with $m_1 = m_2 = 1$ was proposed in [15] under the term double differential PSK (DDPSK). Probability of error was also derived for symbol by symbol detection [12], [15]. Binary and M-ary PSK and the possibility of multiple symbol detection was discussed in [18]. Cascading n first-order differential encoding blocks and numerically computable probability of error expressions for binary DDPSK were reported in [19]. Approximately 4 dB of excess SNR is required to attain a given error probability using binary DDPSK relative to binary DPSK without Doppler frequency.

Here, we exploit the degrees of freedom offered by the different lags in order to improve system performance with respect to the DDPSK proposed in [12] and [15] and later in [18] and [19]. Unfortunately, the exact evaluation of the error probability performance of the system for arbitrary lags in the presence of Doppler frequency and Doppler rate is difficult if not impossible. To select the lags, we adopt the maximum deflection (MD) criterion (or maximum output signal-to-noise power ratio) [13].

IV. Maximum Deflection for Selecting Lags

Consider the case of M-ary PSK signaling (as in Section II-C), and suppose we want to remove a constant phase θ_0 (let us assume $f_e = \alpha_e = 0$).

A. Deflection of $x_2(n; m_1)$

The encoding strategy given by (15) forms the recursion $w_d(n) = w(n)w_d(n-m_1)$. Correspondingly, at the receiver, the decision is based on the noisy estimate of $w(n)$, that is

$$\begin{aligned} x_2(n, m_1) = x(n)x^*(n-1) &= [w_d(n)e^{j\theta_0} + v(n)] \\ &\quad \cdot [w_d^*(n-m_1)e^{-j\theta_0} + v^*(n-m_1)] \\ &= w(n) + w_d(n)e^{j\theta_0}v^*(n-m_1) \\ &\quad + w_d^*(n-m_1)e^{-j\theta_0}v(n) + v(n)v^*(n-m_1) \\ &:= w(n) + d(n; m_1) \end{aligned} \tag{19}$$

where we denoted by $d(n; m_1)$ the disturbance term that is superimposed to the useful term $w(n)$. Under the assumptions that $w(n)$ belongs to an MPSK constellation and $v(n)$ is a zero mean white Gaussian process with variance σ_v^2, both $w(n)$ and $d(n; m_1)$ are zero mean. Thus, we should adopt a definition of deflection that is slightly different from the one in [13]; we define as deflection D the variance ratio of the useful term over the disturbance term

$$D := \frac{\text{var}\{w(n)\}}{\text{var}\{d(n; m_1)\}} = \frac{E\{|w(n)|^2\}}{E\{|d(n; m_1)|^2\}}. \tag{20}$$

The goal is to find the lags that maximize the deflection. The m_1 that maximizes D may not always minimize the probability of error, but we expect it to improve performance relative to $m_1 = 1$. Inserting (19) in (20) and observing that $|w(n)| = |w_d(n)| = 1$, we easily obtain $D = 1/(2\sigma_v^2 + \sigma_v^4)$, which does not depend on m_1; hence, our best choice is $m_1 = 1$ because it minimizes the number of "training" symbols required to initialize the differential decoder [3, p. 211].

However, what if the noise samples are correlated due to a mismatch between $g_c^{(\text{rec})}(t)$ and $g_c^{(\text{tr})}(t)$? To investigate this point further, we suppose that because of filtering effects $v(n)$ is an MA(1) process; it can thus be modeled as the output of a first-order FIR filter with a white noise process at the input, i.e., $v(n) = v_I(n) + bv_I(n-1)$, where $v_I(n)$ is an i.i.d. zero mean Gaussian sequence with variance σ_I^2 such that $\sigma_v^2 = (1+b^2)\sigma_I^2$. The deflection now becomes

$$\begin{aligned} D(m_1) = \Bigg[& 2\sigma_v^2 + \sigma_v^4 + \frac{(1+b^4)\sigma_v^4}{(1+b^2)^2}\delta(m_1) \\ & + \frac{b^2\sigma_v^4}{(1+b^2)^2}[\delta(m_1-1) + \delta(m_1+1)] \Bigg]^{-1} \end{aligned} \tag{21}$$

and depends on m_1. Causality requires $m_1 > 0$, and if deflection is to be maximized with the minimum number of redundant symbols, the best choice is $m_1 = 2$. In general, if L_v denotes the memory of $v(n)$, the minimum m_1 that maximizes deflection is $m_1 = L_v + 1$. Thus, when the noise samples are correlated, we can improve the performance of the differential encoding system by simply selecting a lag different from $m_1 = 1$ that is adopted in the standard DPSK system, [3, ch. 5].

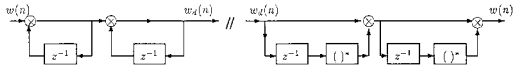


Fig. 3. Second-order differential encoder/decoder $(m_1 = m_2 = 1)$.

The advantage offered by multiple lags is even more pronounced when Doppler frequency is also present, and the decoder implements the third-order ml-HIM to pre-compensate for it.

B. Deflection of $x_3(n; m_1, m_2)$

The encoder output is

$$w_d(n) = w(n)w_d(n-m_1)w_d(n-m_2) \times w_d^*(n-m_1-m_2) \tag{22}$$

and the noisy estimate of $w(n)$ is given by

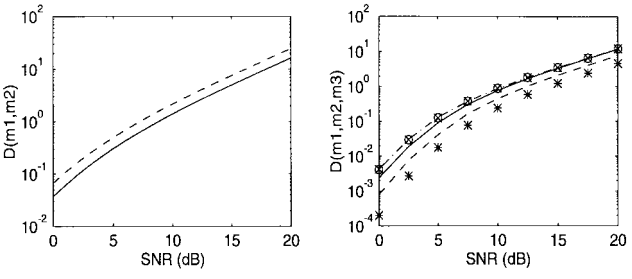
$$\begin{aligned} x_3(n;m_1,m_2) &= x(n)x^*(n-m_1)x^*(n-m_2)x(n-m_1-m_2) \\ &= [w_d(n)e^{j[\theta_0+2\pi f_e n]} + v(n)][w_d^*(n-m_1) \\ &\quad \times e^{-j[\theta_0+2\pi f_e(n-m_1)]} + v^*(n-m_1)] \\ &\quad \times [w_d^*(n-m_2)e^{-j[\theta_0+2\pi f_e(n-m_2)]} + v^*(n-m_2)] \\ &\quad \times [w_d(n-m_1-m_2)e^{j[\theta_0+2\pi f_e(n-m_1-m_2)]} \\ &\quad + v(n-m_1-m_2)]. \end{aligned} \tag{23}$$

Carrying out all the products in (23), we obtain the result in the form $x_3(n; m_1, m_2) = w(n) + d(n; m_1, m_2)$, where the disturbance consists of 15 terms. Evaluation of $E\{|d(n; m_1, m_2)|^2\}$ involves 15×15terms, but most of them do not contribute to the final result because they have zero mean. Skipping details, we find

$$D(m_1,m_2) = \left[4\sigma_v^2 + 6\sigma_v^4 + 4\sigma_v^6 + \sigma_v^8 + \left(2\sigma_v^2 + 5\sigma_v^4 + 4\sigma_v^6 + \sigma_v^8\right)\delta(m_1 - m_2)\right]^{-1}. \tag{24}$$

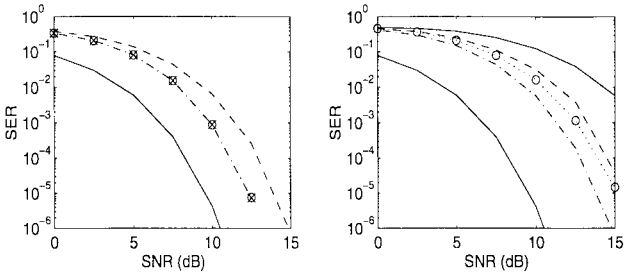
The result in (24) was verified with Monte-Carlo simulations. The deflection in (24) does not depend on the values of θ_0, f_e, and α_e. The choice $m_1 = m_2 = 1$ corresponds to the DDPSK system proposed in [15] and [19] and is the worst one in terms of SNR. On the contrary, $m_2 > m_1 = 1$ offers the same performance, and in order to minimize the training symbols, we select $m_2 = 2$. Fig. 4(a) shows $D(m_1, m_2)$ in (24) versus SNR with the solid line corresponding to $m_1 = m_2$ and the dashed to $m_1 \neq m_2$. In Fig. 4(a), plots of $D(m_1, m_2)$ in (24) are reported versus the signal-to-noise ratio, which is defined as SNR $:= E\{|w(n)|^2\}/\sigma_v^2$. The solid line corresponds to $m_1 = m_2$, whereas the dashed line refers to $m_1 \neq m_2$.

The deflection of $x_4(n; m_1, m_2, m_3)$ can be evaluated along the same lines, but the calculation involves $(2^8 - 1) \times (2^8 - 1)$ terms. We derived it via Monte Carlo simulations using 5×10^5 symbols for each set of lags. Fig. 4(b) depicts $D(m_1, m_2, m_3)$ versus SNR for $(m_1, m_2, m_3) = (1,1,1)$, $(1,1,2)$, $(1,2,4)$, $(2,3,4)$, $(1,2,3)$, and $(1,4,8)$. Cascading three first-order differential encoding (decoding) cells corresponds to $(m_1, m_2, m_3) = (1,1,1)$ and yields the lowest deflection due to the presence of dependent samples in $x_4(n; m_1, m_2, m_3)$ (note that samples $x(n-1)$ and $x(n-2)$ appear three times each in x_4). When the number of dependent samples in the ml-HIM decreases, the deflection increases. When $(m_1, m_2, m_3) = (1,1,2)$, x_4 contains $x(n)$, $x(n-4)$, and two replicas of $x(n-1)$, $x(n-2)$, and $x(n-3)$. When $(m_1, m_2, m_3) = (1,2,3)$, $x(n-3)$ is present two times; in all other cases, the noise samples are independent, and the deflection reaches its maximum.



(a) (b)

Fig. 4. Deflection versus SNR of (a) third-order ml-HIM for $m_1 = m_2$ (solid line) and $m_1 < m_2$ (dashed line). (b) Fourth-order ml-HIM for $(m_1, m_2, m_3) = (1,1,1)$ (stars), $(1,1,2)$ (dashed line), $(1,2,4)$ (dash-dotted line), $(2,3,4)$ (circles), $(1,2,3)$ (solid line), and $(1,4,8)$ (x-marks).



(a) (b)

Fig. 5. Binary PSK. SER versus SNR with (a) third-order ml-HIM for $(m_1, m_2) = (1,1)$ (dashed line), $(1,2)$ (dash-dotted), $(1,3)$ (circles), and $(1,4)$ (x-marks). (b) Fourth-order ml-HIM for $(m_1, m_2, m_3) = (1,1,1)$ (solid line), $(1,1,2)$ (dashed), $(1,2,4)$ (dotted), $(2,3,4)$ (circles), and $(1,2,3)$ (dash-dotted). The lower solid lines in (a) and (b) denote the SER of an ideal BPSK $(\theta_e(t) = 0)$.

Symbol error rate (SER) curves were obtained by averaging over 8×10^7 decisions for binary PSK with parameters $\theta_0 = \pi/8$, $f_e = 0.05$, $\alpha_e = 0$ [Fig. 5(a)], and $\alpha_e = 0.0015$ [Fig. 5(b)]. Fig. 5(a) shows the SER for lags $(m_1, m_2) = (1,1)$ (dashed line) $(1,2)$ (dash-dotted), $(1,3)$ (circles), and $(1,4)$ (x-marks). The solid lines in Figs. 5(a) and (b) denote the SER of an ideal (i.e., with $\theta(n) = 0$) binary PSK. The lags $(m_1, m_2) = (1,1)$ yield the same performance as in the DDPSK system of [19]. As expected, the choice $(m_1, m_2) = (1,1)$ is the worst one also in terms of probability of error performance, whereas all other choices are equivalent; hence, $(m_1, m_2) = (1,2)$ should be selected because it minimizes the number of training symbols. For a target SER of 10^{-5}, 4.5 dB of additional SNR is required when $m_1 = m_2 = 1$ are used instead of ideal BPSK with $\theta_0 = f_e = 0$. However, if we select $(m_1, m_2) = (1,2)$, approximately 2.8 dB of additional SNR is only necessary. It is worth observing that the SER of this

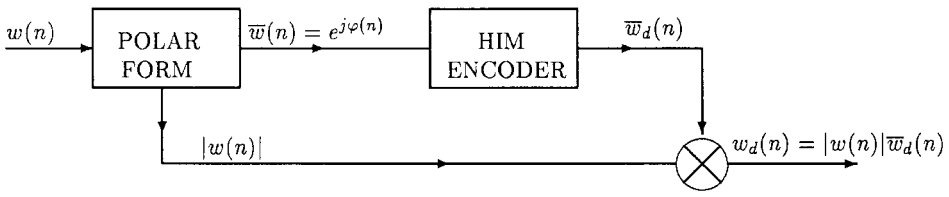


Fig. 6. Differential encoder for nonconstant modulus constellations.

system does not depend[2] on the values of θ_0 and f_e. Obviously, it would be impossible to compensate for a nonzero α_e.

To remove Doppler rate effects, the system has to implement the fourth-order ml-HIM at the decoder. Fig. 5(b) shows the SER of such a system for $\theta_0 = \pi/8$, $f_e = 0.05$, and $\alpha_e = 0.0015$ with lags $(m_1, m_2, m_3) = (1,1,1)$ (upper solid line), $(1,1,2)$ (dashed), $(1,2,4)$ (dotted), $(2,3,4)$ (circles), and $(1,2,3)$ (dash-dotted). The results show that selecting the lags appropriately improves the performance considerably, relative to the system in [19], which uses $m_1 = m_2 = m_3 = 1$. For example, selecting $(m_1, m_2, m_3) = (1,2,3)$ requires only 4.5 dB of additional SNR relative to the ideal BPSK, whereas adopting $m_1 = m_2 = m_3 = 1$ incurs greater loss. We note that all sets of lags that correspond to independent noise samples guarantee the same SER, but the optimal choice in terms of SER is $(m_1, m_2, m_3) = (1,2,3)$, whose deflection is slightly lower than that corresponding to the maximum D [see Fig. 4(b)].

V. Nonconstant Modulus Constellations

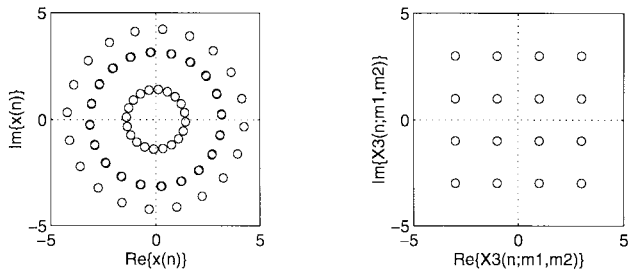
Thus far, we have considered only M-ary PSK constellations, and interestingly, in the technical literature, differential encoding has been restricted only to MPSK modulation systems, i.e., to constant modulus constellations. Relying on the ml-HIM approach, it is shown next how generalized differential encoding techniques can be successfully applied to generic (nonconstant modulus) constellations, such as the QAM constellations and amplitude-modulation phase-modulation (AM-PM) [3, ch. 5].

A. 16-QAM Modulations

Consider a 16-QAM modulation scheme. The input $w(n) = a(n) + jb(n) = \rho(n)\exp[j\varphi(n)]$ is assumed to be i.i.d., with $a(n)$ and $b(n)$ drawn from the discrete alphabet set $\{\pm 1, \pm 3\}$. Suppose that only phase and Doppler frequency errors are present in $x(n)$. To remove phase and Doppler frequency errors with the third-order ml-HIM encoder, we define $\bar{w}(n) := w(n)/\rho(n) = \exp[j\varphi(n)]$, and $\bar{w}_d(n) := w_d(n)/\rho(n)$. At the differential encoder output, we have (see Fig. 6) $w_d(n) = \rho(n)\bar{w}_d(n)$, where, according to (13), $\bar{w}_d(n)$ is

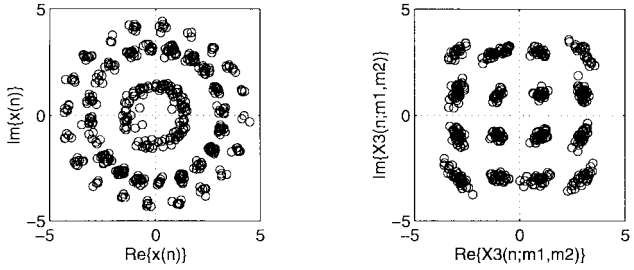
$$\bar{w}_d(n) = \bar{w}(n)\bar{w}_d(n-m_1)\bar{w}_d(n-m_2)\bar{w}_d^*(n-m_1-m_2). \tag{25}$$

[2]This is only approximately true, as far as $\lambda(f_e, \alpha_e) \approx 1$; otherwise, we should consider the effective signal-to-noise ratio as $\text{SNR}_e = |\lambda(f_e, \alpha_e)|^2\text{SNR}$.



(a) (b)

Fig. 7. Sixteen-QAM. (a) Received uncoded signal (SNR $= \infty$). (b) Third-order ml-HIM (SNR $= \infty$).



(a) (b)

Fig. 8. Sixteen-QAM. (a) Received uncoded signal (SNR $= 25$ dB). (b) Third-order ml-HIM (SNR $= 25$ dB).

The information sequence is then recovered by properly normalizing the ml-HIM of the received signal

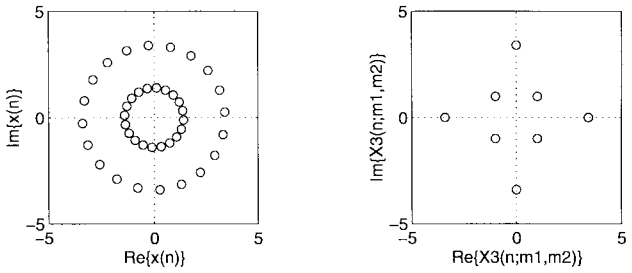
$$\bar{x}_3(n; m_1, m_2) = \frac{x_3(n; m_1, m_2)}{|x(n-m_1)x(n-m_2)x(n-m_1-m_2)|} \tag{26}$$

and applying the decision rule in (18) as

$$\hat{w}(n) = \arg\min_{w(n)} |w(n) - \bar{x}_3(n; m_1, m_2)|. \tag{27}$$

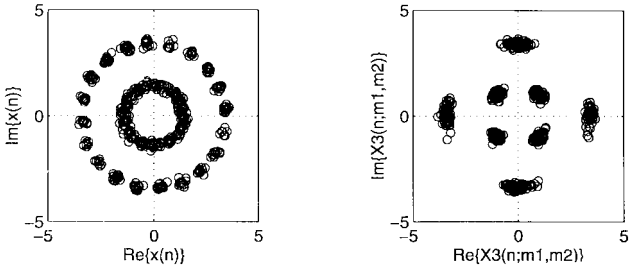
The normalization is necessary to remove amplitude effects of the previous transmitted symbols since only their phases are necessary for differential decoding. In Fig. 7(a), we plot the real versus imaginary parts of the received signal (2000 noise-free 16-QAM symbols) when the uncoded $w(n)$ is transmitted over the channel with phase and frequency errors ($\theta_0 = \pi/8$, $f_e = 0.05$, and $\alpha_e = 0$). Fig. 7(b) depicts real versus imaginary parts of $\bar{x}_3$ when $w_d(n)$ is transmitted with differentially encoded phase according to (25). It is evident that in the absence of noise, the constellation is recovered perfectly. The same plots with noise at SNR $= 25$ dB are depicted in Fig. 8(a) and (b).

Note that the encoder block diagram shown in Fig. 6 is quite general. It is valid for every linear modulation format, as



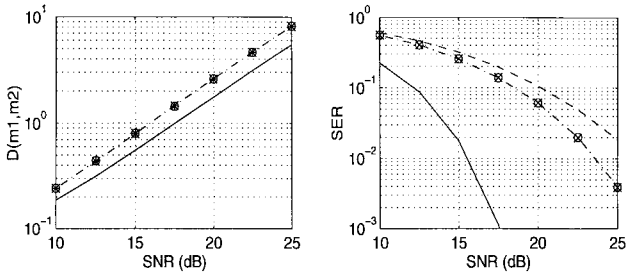
(a) (b)

Fig. 9. Eight AM-PM. (a) Received uncoded signal (SNR $= \infty$). (b) Third-order ml-HIM (SNR $= \infty$).



(a) (b)

Fig. 10. Eight AM-PM. (a) Received uncoded signal (SNR $= 25$ dB). (b) Third-order ml-HIM (SNR $= 25$ dB).



(a) (b)

Fig. 11. Sixteen-QAM. (a) Deflection of third-order ml-HIM versus SNR for $(m_1, m_2) = (1,1)$ (solid line), $(1,2)$ (dash-dotted line), $(1,3)$ (circles), $(1,4)$ (stars). (b) SER versus SNR with third-order ml-HIM decoder for $(m_1, m_2) = (1,1)$ (dashed line), $(1,2)$ (dash-dotted line), $(1,3)$ (circles), and $(1,4)$ (x-marks). The solid line in (b) denotes the SER of an ideal 16-QAM $(\theta_c(t) = 0)$.

far as the "HIM encoder" block implements the relationships (15)–(17) and the decision strategy (26)–(27) is applied, with the appropriate order (higher than three if necessary). As an example, Figs. 9 and 10 show plots similar to Figs. 7 and 8 with the same set of parameters but for an eight AM-PM constellation, which is a combined modulation with two amplitudes $(r_1 = \sqrt{2},\ r_2 = 2 + \sqrt{2})$ and four phases for each amplitude [3, p. 221]. Altogether, Figs. 7–10 show that frequency errors, even in the absence of Doppler rate, deteriorate system performance significantly, if a proper differential encoding/decoding is not applied.

Fig. 11(a) and (b) show the deflection of $\bar{x}_3$ and the corresponding SER versus SNR for $(m_1, m_2) = (1,1)$, $(1,2)$, $(1,3)$, and $(1,4)$ obtained based on 2×10^5 symbols and with parameters $\theta_0 = \pi/8$, $f_e = 0.05$, and $\alpha_e = 0$. Due to the multilevel constellation (which is less tolerant to noise than BPSK), the loss is greater than in Fig. 3(b); however, part of this loss can be recovered if the symbol-by-symbol decision is replaced by multiple symbol detection, as suggested in [18] (see also [19]). The idea of multiple symbol differential detection was recently extended also to Rayleigh fading channels [11]. These directions are beyond the scope of this work and will be pursued in future works.

VI. Conclusion

The idea of differential encoding has been reinterpreted, generalized to higher orders, and extended to nonconstant modulus signals relying on the ml-HIM.

Specifically, the input–output relationships of the well-known DPSK system (first-order differential encoding system for PSK signals) and the less known DDPSK system (second-order DPSK) have been reinterpreted using the ml-HIM formulation. In fact, DPSK and DDPSK turned out to be particular cases of a generalized differential decoder corresponding to the second- and third-order HIM for a particular choice of the lags. Exploiting the degrees of freedom offered by the different lags, we can increase considerably the system performance without increasing complexity. The maximum deflection criterion was used as a tool for selecting the lags. If necessary, higher-than-third-order ml-HIM can be implemented at the decoder to precompensate for higher order phase distortions. Within the ml-HIM framework, generalized differential encoding was also derived for nonconstant modulus constellations, such M-ary QAM and AM-PM.

In this work, we have not considered the presence of ISI. When this is not a reasonable assumption (e.g., in frequency-selective fading channels), the encoding/decoding system should be considered to be a complement to blind equalization. Note that even if the channel is linear, its cascade with the nonlinear (ml-HIM based) decoder at the receiver gives rise to an overall nonlinear transformation of the transmitted symbols. Mitigation of such a nonlinear effect with linear equalizers is expected to perform rather poorly, which suggests, as a possible future research direction, usage of nonlinear equalizers. A linear equalizer can be used before the ml-HIM decoder, but this would exclude the possibility of using a decision feedback structure. Alternatively, the use of a fractionally spaced nonlinear equalizer at the output of the differential detector was proposed in [7]. Only a first-order differential detector was considered in [7]; the extension of this approach to generalized differential encoding systems seems to be an interesting topic for future research.

References

[1] S. Barbarossa, "Detection and estimation of the instantaneous frequency of polynomial phase signals by multilinear time-frequency representations," in *Proc. IEEE-SP Workshop Higher Order Statist.*, Lake Tahoe, CA, June 1993, pp. 168–172.

[2] S. Barbarossa, A. Scaglione, and G. B. Giannakis, "Product multilag high-order ambiguity function for multicomponent polynomial phase signal modeling," *IEEE Trans. Signal Processing*, vol. 46, pp. 691–708, Mar. 1998.

[3] S. Benedetto, E. Biglieri, and V. Castellani, *Digital Transmission Theory*. Englewood Cliffs: Prentice-Hall, 1987.

[4] F. Gini, M. Luise, and R. Reggiannini, "Analysis and design of a DPSK optical heterodyne receiver in the presence of laser phase noise and frequency detuning," *Int. J. Commun. Systems*. New York: Wiley, 1995, vol. 8, pp. 129–141.
[5] L. Chen, H. Kusaka, and M. Kominami, "Blind phase recovery in QAM communication systems using higher order statistics," *IEEE Signal Processing Lett.*, vol. 3, pp. 147–149, May 1996.
[6] M. Luise and R. Reggiannini, "Carrier frequency acquisition and tracking for OFDM systems," *IEEE Trans. Commun.*, vol. 44, pp. 1590–1598, Nov. 1996.
[7] A. Masoomzadeh-Fard and S. Pasupathy, "Nonlinear equalization of multipath fading channels with noncoherent demodulation," *IEEE J. Select. Areas Commun.*, vol. 14, pp. 512–520, Apr. 1996.
[8] M. McIntyre and A. Ashley, "A simple fixed-lag algorithm for tracking frequency rate-of-change," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, pp. 677–683, July 1993.
[9] S. Kay, "A fast and accurate single frequency estimator," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 1987–1990, Dec. 1989.
[10] A. Kajiwara, "Mobile satellite CDMA system robust to doppler shift," *IEEE Trans. Veh. Technol.*, vol. 44, pp. 480–486, Aug. 1995.
[11] S. L. Miller and R. J. O'Dea, "Multiple symbol differential detection for trellis-coded MPSK over Rayleigh fading channels," *IEEE Commun. Lett.*, vol. 1, pp. 2–4, Jan. 1997.
[12] Y. B. Okunev, *Theory of Phase-Difference Modulation*. Moscow, U.S.S.R.: Svyaz, 1979.
[13] B. Picinbono, "On deflection as a performance criterion in detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, pp. 1072–1081, 1995.
[14] S. Peleg and B. Porat, "Estimation and classification of signals with polynomial phase," *IEEE Trans. Inform. Theory*, vol. 37, pp. 422–430, Mar. 1991.
[15] M. Pent, "Doubly differential PSK scheme in the presence of Doppler shift," in *Proc. Digital Commun. Avionics, AGARD*, 1978, pp. 43.1–43.11.
[16] B. Porat, *Digital Processing of Random Signals, Theory and Methods*. Englewood Cliffs, NJ: Prentice-Hall, 1994.
[17] A. W. Rihaczek, *Principles of High-Resolution Radar*. San Diego, CA: Peninsula, 1985.
[18] M. K. Simon and D. Divsalar, "On the implementation and performance of single and double differential detection schemes," *IEEE Trans. Commun.*, vol. 40, pp. 278–291, Feb. 1992.
[19] D. K. van Alphen and W. C. Lindsey, "Higher-order differential phase shift keyed modulation," *IEEE Trans. Commun.*, vol. 42, pp. 440–448, Feb./Mar./Apr. 1994.
[20] G. Vannucci and G. J. Foschini, "Characterizing filtered light waves corrupted by phase noise," *IEEE Trans. Inform. Theory*, vol. 34, pp 1437–1448, Nov. 1988.
[21] B. Yu, Y. B. Okunev, and L. M. Fink, "Noise immunity of various receiving methods for binary systems with second-order phase-difference modulation," *Telecommun. Radio Eng.*, vol. 39, no. 8, pp. 51–56, 1984.
[22] G. T. Zhou, G. B. Giannakis, and A. Swami, "On polynomial phase signals with time-varying amplitudes," *IEEE Trans. Signal Processing*, vol. 44, pp. 848–861, Apr. 1996.

Fulvio Gini was born in Fucecchio, Italy, on February 25, 1965. He received the Doctor Engineer (cum laude) and the Research Doctor degrees in electronic engineering from the University of Pisa, Pisa, Italy, in 1990 and 1995 respectively.

He is currently a Tenured Research Fellow with the Dipartimento di Ingegneria dell'Informazione, University of Pisa. From July 1996 through January 1997, he was a Visiting Researcher at the Department of Electrical Engineering, University of Virginia, Charlottesville. His general interests are in the areas of statistical signal processing, estimation, and detection theory. In particular, his research interests include non-Gaussian signal detection and estimation using higher order statistics, cyclostationary signal analysis, and estimation of nonstationary signals with applications to radar and digital communications.

Georgios B. Giannakis (F'96) received the Diploma in electrical engineering from the National Technical University of Athens, Athens, Greece, in 1981. From September 1982 to July 1986, he was with the University of Southern California (USC), where he received the M.Sc. degree in electrical engineering in 1983, the M.Sc. degree in mathematics in 1986, and the Ph.D. degree in electrical engineering in 1986.

After lecturing for one year at USC, he joined the University of Virginia, Charlottesville, in September 1987, where he is now a Professor in the Department of Electrical Engineering. His general interests lie in the areas of signal processing, communications, estimation and detection theory, and system identification. Specific research areas of current interest include diversity techniques for channel estimation and multiuser communications, nonstationary and cyclostationary signal analysis, wavelets in statistical signal processing, and non-Gaussian signal processing with applications to SAR, array, and image processing.

Dr. Giannakis received the IEEE Signal Processing Society's 1992 Paper Award in the Statistical Signal and Array Processing (SSAP) area. He co-organized the 1993 IEEE Signal Processing Workshop on Higher Order Statistics, the 1996 IEEE Workshop on Statistical Signal and Array Processing, and the first IEEE Signal Processing Workshop on Wireless Communications in 1997. He guest (co-)edited two special issues on high-order statistics (*International Journal of Adaptive Control and Signal Processing*, and the EURASIP journal *Signal Processing*) and the January 1997 Special Issue on Signal Processing for Advanced Communications (IEEE TRANSACTIONS ON SIGNAL PROCESSING). He has served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE SIGNAL PROCESSING LETTERS, a secretary of the Signal Processing Conference Board, a member of the SP Publications board and the SSAP Technical Committee, and a chair of the Signal Processing for Communications Technical Committee. He is also a member of the IMS and the European Association for Signal Processing.