Research paper

Low-complexity factor-graph-based MAP detector for filter bank multicarrier systems

CUI Fangyu[1], ZHAO Minjian[1], CAI Yunlong[1], CHAMPAGNE Benoit[2]

1. College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China
2. Department of Electrical and Computer Engineering, McGill University, Montreal H3A0E9, Canada

Abstract: FBMC (Filter Bank Multicarrier) modulation is considered one of the waveform candidates in fifth generation wireless communication technology because of its several improved features compared to conventional orthogonal frequency division multiplexing schemes. A soft-input-soft-output factor-graph-based maximum-a-posterior detector is applied to FBMC systems. The detector achieves better performance than simple linear equalizers such as minimum mean square error and zero forcing in coded systems while exhibiting only a linear growth in complexity with the number of simultaneous interfering symbols. Furthermore, the proposed detector can be easily extended to cases where FBMC modulation is combined with multiple-input-multiple-output processing. The complexity of the detector is analyzed and the simulation results demonstrated its superior performance.

Key words: FBMC, MIMO, factor graph, turbo detection, MAP

1 Introduction

FBMC (Filter Bank Multicarrier) modulation is one of the waveform candidates in 5G (Fifth-Generation) systems because of its several advantages over conventional OFDM (Orthogonal Frequency Division Multiplexing) schemes[1]. First, without a CP (Cyclic Prefix), FBMC modulation can achieve higher spectral efficiency. In addition, it uses well localized prototype filters to reduce the spectral sidelobes, which makes it less sensitive to the interference caused by imperfect synchronization.

This makes FBMC modulation particularly attractive for asynchronous transmissions, as occurs in D2D (Device-to-Device) communication, which is an important use case of 5G systems. Furthermore, the small sidelobe makes it well-suited for cognitive radio applications.

FBMC modulation is non-orthogonal and introduces an imaginary interference which can be eliminated by acquiring a real one. However, when the subchannels are non-flat, the introduced interference will not be purely imaginary and ISI (Inter-Symbol Interference) as well as ICI (Inter-Carrier Interference) will

arise. Usually, the ZF (Zero Forcing) or MMSE (Minimum Mean Square Error) criteria combined with per subcarrier channel equalizers are considered to deal with the interference because of their low computational complexity[2]. When the channel is perfectly equalized, the interference will be restored to the purely imaginary domain and will not disturb the detection. If FBMC modulation is combined with MIMO (Multiple-Input Multiple-Output) systems, MMSE and ZF equalizers will equalize the IAI (Inter-Antenna Interference) of each frequency point at the same time[3].

Although linear equalizers have simple structures, their performance is limited. To increase the detection performance, one needs to consider the optimal ML (Maximum-Likelihood) and MAP (Maximum-A-Posterior) criteria[4]. Unfortunately, owing to the 2D (Two-Dimensional) nature of the interference in FBMC systems (i.e. time and frequency), the complexity of traditional Viterbi and BCJR algorithms will grow exponentially with the length of ISI and the number of subcarriers. In MIMO-FBMC systems, there will be an additional IAI along the spatial dimension, which makes the complexity grow exponentially with the number of transmitting antennas. This motivates the use of low-complexity, suboptimal ML and MAP detectors with good performance.

Applying the SPA (Sum-Product Algorithm) to a FG (Factor Graph) is a method for achieving the ML and MAP criteria[5]. To reduce the complexity, one can employ an FG with cycles, although it cannot provide an exact convergence. However, when the shortest cycle is longer than six, the SPA can converge approximately and provide a good approximation to the optimal ML/MAP solution under practical conditions[5]. In this paper, an FG-based MAP detector[6] is extended to FBMC and MIMO-FBMC systems. The FG is derived from the Ungerboeck model[7] without any approximation; its

shortest cycle is longer than six, and accordingly, a good performance can be achieved[6,8]. Because of the FG structure, the computational complexity of the detector grows linearly with the number of interferers, defined here as the number of neighboring symbols in the time-frequency lattice that interfere with the detected symbol. When we extend the detector to MIMO-FBMC systems, so that it may handle IAI, its complexity will also grow linearly with the number of antennas. The SISO (Soft-Input-Soft-Output) property of the detector makes it suitable to be combined with a turbo structure to improve the performance of coded systems, which is a key advantage over MMSE and ML equalizers[9,10].

The rest of the paper is organized as follows. The system model of FBMC modulation is described in Section 2 and the FG-based SISO detector is derived in Section 3. The extension of the detector to MIMO-FBMC is considered in Section 4. The complexity analysis and numerical simulation of the detector are discussed in Section 5 and Section 6, respectively. Finally, the conclusions are presented in Section 7.

2 System model

FBMC modulation is a method that employs filter banks to obtain narrow subcarriers to reduce the overlap between different subchannels. Several types of FBMC modulation methods have been proposed in the literature, such as: SMT (Staggered Multitone), CMT (Cosine-Modulated Multitone) and FMT (Filtered Multitone)[1]. SMT is based on OQAM (Offset Quadrature Amplitude Modulation) and it is currently attracting much interest in the literature owing to its typicality; Accordingly, SMT is considered as a specific type of FBMC modulation in this paper.

After SMT modulation, the transmitted baseband signal can be expressed as

$$s(m) = \sum_{k=1}^{M} \sum_{n=1}^{N} d_{k,n} \theta_{k,n} h(m - nM/2) e^{jk\frac{2\pi m}{M}}, \qquad (1)$$

where k, n are the subcarrier and symbol index, respectively, $d_{k,n}$ is the transmitted symbol on subcarrier k at symbol interval n, h is the prototype filter, M is the number of subcarriers and N is the number of symbols in a data block. The multiplicative parameter $\theta_{k,n}$ in Eq.(1), which serves to form the OQAM structure, is defined as

$$\theta_{k,n} \begin{cases} \pm 1, & \text{for } n+k \text{ even}, \\ \pm j, & \text{for } n+k \text{ odd}. \end{cases} \qquad (2)$$

The received baseband signal is affected by multipath (frequency selective) radio propagation and additive white Gaussian noise. It can be expressed as

$$r(m) = s(m) * c(m) + v(m)$$
$$= \sum_{k=1}^{M} \sum_{n=1}^{N} d_{k,n} \theta_{k,n} (h(m - nM/2) e^{jk\frac{2\pi m}{M}} * c(m)) + v(m),$$
$$\qquad (3)$$

where $*$ denotes the discrete-time convolution, $c(m)$ is the impulse response of the propagation channel and $v(m)$ is the zero-mean noise term with variance N_0.

There are two types of SMT receiver structures: PPN (Polyphase Network)[2] and FS (Frequency Despreading)[11,12]. PPN needs an M-point FFT (Fast Fourier Transform) operation, whereas FS needs an FFT operation whose size must equal the length of the prototype filter to cover the full spectrum. Hence, the PPN structure is frequently used in the literature because of its lower complexity. However, FS naturally allows a straightforward multipoint equalization in the frequency domain with no delay[11], which makes it more suitable for use in cases where the subchannels are non-flat fading.

3 FG-based detector

3.1 PDF reorganization

To construct the FG for the proposed detector, we require the PDF (Probability Density Function) of the complete received sequence to be properly reorganized and then factorized into a product of individual PDFs for each one of the variables. After a proper PDF reorganization, we can formulate a factor graph with its shortest cycles longer than six.

The FG-based SISO detector is derived using the Ungerboeck model as described in Ref.[6]. From Eq.(3), the conditional PDF obeys the following relationship,

$$p(\boldsymbol{s} \mid \boldsymbol{r}) \propto \exp\left(-\frac{\sum_m |r(m) - s(m) * c(m)|^2}{2N_0} \right)$$
$$= \exp\left[-\frac{1}{2N_0}\left(\sum_m |r(m)|^2 + \sum_m |s(m) * c(m)|^2 \right. \right.$$
$$\left. \left. -2\text{Re}\left\{ \sum_m r(m)(s(m) * c(m))^* \right\} \right) \right], \qquad (4)$$

where Re denotes the real part. The argument of the exponential function in Eq.(4) consists of a sum of three terms. The first term $\sum_m |r(m)|^2$ is constant and can be ignored, whereas the remaining terms are discussed in more detail below.

The second term in Eq.(4) can be expanded as

$$\sum_m |s(m) * c(m)|^2 = \sum_{k_1,k_2=1}^{M} \sum_{n_1,n_2=1}^{N} d_{k_1,n_1}^* \theta_{k_1,n_1}^*$$
$$\cdot d_{k_2,n_2} \theta_{k_2,n_2} \lambda(k_1, k_2, n_1, n_2), \quad (5)$$

where

$$\lambda(k_1, k_2, n_1, n_2) = \sum_m \left[\left(h(m - n_1 M/2) e^{jk_1\frac{2\pi m}{M}} * c(m) \right)^* \right.$$
$$\left. \cdot \left(h(m - n_2 M/2) e^{jk_2\frac{2\pi m}{M}} * c(m) \right) \right]. \quad (6)$$

From Eq.(6), it is clear that $\lambda = \lambda(k_1, k_2, n_1, n_2)$ satisfies the conjugate symmetry property

$$\lambda(k_1, k_2, n_1, n_2) = \lambda^*(k_2, k_1, n_2, n_1). \qquad (7)$$

Here $\lambda(k_1, k_2, n_1, n_2)$ can be regarded as a coefficient characterizing the interference between d_{k_1,n_1} and d_{k_2,n_2}. Because of the localization property of the prototype filter, the symbols that can create a significant

interference with $d_{k,n}$ are those nearly located in the time-frequency plane. Accordingly, we make the following assumption:

$$\lambda(k_1,k_2,n_1,n_2)=0, \text{ for } |k_1-k_2|>L_1 \text{ or } |n_1-n_2|>L_2, \text{ (8)}$$

where L_1 and L_2 are the range of frequency and time that define the maximum distances between two symbols that can interfere with each other. The values of these parameters are mainly affected by the frequency and time width of the prototype filter. Based on Eq.(8), the maximum number of neighboring symbols that interfere with each other, i.e., the upper bound of the number of interferers, denoted by L, will be limited to $(2L_1+1)(2L_2+1)-1$. In practice, we find that L can be further reduced as many of the terms $\lambda(k_1, k_2, n_1, n_2)$ within a given time-frequency rectangle as in Eq.(8) are nearly zero and can be ignored.

The third term in Eq.(4) can be expanded as

$$2\text{Re}\left\{\sum_m r(m)(s(m)*c(m))^*\right\}=2\text{Re}\left\{\sum_{k=1}^{M}\sum_{n=1}^{N}d_{k,n}^*\theta_{k,n}^* y_{k,n}\right\}$$
(9)

where

$$y_{k,n}=\sum_m r(m)(h(m-nM/2)e^{jk\frac{2\pi m}{M}}*c(m))^* \quad (10)$$

can be seen as the received signal $r(m)$ correlated with the filter $\left(h(m-nM/2)e^{jk\frac{2\pi m}{M}}*c(m)\right)^*$, i.e., the matched filter corresponding to the combination of transmit filter and channel response.

Utilizing the two properties Eqs.(7) and (8) of λ, we can write as

$$p(\boldsymbol{r}\mid\boldsymbol{s})\propto\exp\left[\frac{1}{N_0}\left(\text{Re}\left\{\sum_m r(m)(s(m)*c(m))^*\right\}\right.\right.$$
$$\left.\left.-\frac{1}{2}\sum_m|s(m)*c(m)|^2\right)\right]$$
$$=\exp\left[\frac{1}{N_0}\sum_{k=1}^{M}\sum_{n=1}^{N}\text{Re}\left\{d_{k,n}^*\theta_{k,n}^*\right.\right.$$

$$\cdot\left(y_{k,n}-\frac{1}{2}d_{k,n}\theta_{k,n}\lambda(k,k,n,n)\right.$$
$$-\sum_{k_2=k-L_1}^{k+L_1}\sum_{n_2=n-L_2}^{n-1}d_{k_2,n_2}\theta_{k_2,n_2}\lambda(k,k_2,n,n_2)$$
$$\left.\left.\left.-\sum_{k_2=k-L_1}^{k-1}d_{k_2,n}\theta_{k_2,n}\lambda(k,k_2,n,n)\right)\right\}\right]. \quad (11)$$

3.2 PDF factorization and factor graph representation

Depending on the channel characteristics, the performance and complexity of the FG-based approach will be different. When the sub-channel is approximately flat fading, one point of observation for each sub-channel is enough to estimate the channel influence with enough precision. However, when the sub-channel is highly frequency selective, several observations per sub-channel are needed to calculate the interference between symbols while the complexity is higher. In this paper we assume that K observations per sub-band are sufficient to determine the influence of the channel, where K is the overlap factor of the prototype filter. Equalization with fewer points can be easily derived from this case by simply sampling the frequency points considered in calculating the interference term.

As the overlap factor of the prototype filter is K, the length of the prototype filter will be KM sample points. Accordingly KM is the proper number of points to express the channel at one symbol interval. Thus the channel is expressed using a total of KM coefficients, i.e., an average of K coefficients per sub-band in the frequency domain.

Based on this observation, the input vector of the detector at symbol interval n is

$$\boldsymbol{r}_n=[r(nM-KM/2+1),\cdots,r(nM+KM/2)]^{\text{T}}. \quad (12)$$

After a KM-point FFT operation, the frequency domain vector $\boldsymbol{R}_n=FFT[\boldsymbol{r}_n]$ can be obtained. Similarly, we can define $\boldsymbol{C}_n=FFT[\boldsymbol{c}_n]$ as the frequency domain

channel coefficients at symbol interval *n*, where $c_n = [c_n(0), c_n(1), \cdots c_n(KM)]^T$ is the corresponding time domain channel impulse response, and $\boldsymbol{H} = FFT[\boldsymbol{h}]$ is the frequency domain representation of the prototype filter $\boldsymbol{h} = [h(0), h(1), \cdots, h(KM)]^T$. Because of the frequency localization property of the filter, the points out of the passband in $\boldsymbol{H}$ are zero or very close to zero and can be ignored. The number of nonzero points W is equal to the width of the filter in the frequency domain. Although W is different for different prototype filters, it is always proportional to K. For PHYDYAS filters W is given by $2K-1$[13] and for IOTA filters it is about $7K-1$[14] when the related design parameters are $v_0 = 1$, $\tau_0 = 1/2$, $\rho = 2$.

The matched filtering operation in Eq.(10) can be expressed as

$$y_{k,n} \approx \frac{1}{KM} \sum_{l=0}^{KM-1} H^*(l-kK)C_n^*(l)R_n(l), \quad (13)$$

where $H(l)$ is the lth component of vector $\boldsymbol{H}$, $R_n(l)$ is the lth component of vector $\boldsymbol{R}_n$ and $C_n(l)$ is the lth component of vector $\boldsymbol{C}_n$. Here we use the KM-point circular convolution instead of the linear convolution of two KM-point pulses and thus '$\approx$' is used. Because of the localized structure of filter vector $\boldsymbol{H}$, there are only W non-zero points out of KM-points that need to be calculated. Accordingly, we write

$$y_{k,n} \approx \frac{1}{KM} \sum_{l=kK-(W-1)/2}^{kK+(W-1)/2} H^*(l-kK)C_n^*(l)R_n(l). \quad (14)$$

This filtering operation can be directly implemented by the FS structure as $\boldsymbol{R}_n$ is calculated at the receiver[11]. If we want to use the PPN structure to reduce the complexity, the sub-channel filters $\{C_n^*(l); kK-(W-1)/2 < l < kK+(W-1)/2\}$ have to be converted into a time domain[2].

Similar to Eq.(13), λ can also be expressed in frequency domain as follows

$$\lambda(k_1, k_2, n_1, n_2) \approx \frac{1}{KM} \sum_{l=0}^{KM-1} H^*(l-k_1 K)C_{n_1}^*(l)$$
$$\cdot H(l-k_2 K)C_{n_2}(l)e^{j(n_1-n_2)\frac{M}{2}\frac{2\pi l}{KM}}. \quad (15)$$

Again, most W points are nonzero which need to be calculated. However, because of the localization property of the prototype filter, the approximation is good when the two symbols are close to each other, i.e., the central points of the two pulses are close. When the symbols are far from each other, the correlation of the two pulses is usually small and can be ignored in practical conditions.

Based on the above expressions of $y_{k,n}$ and λ (k_1, k_2, n_1, n_2), the following functions can be defined

$$F_{k,n}(d_{k,n}) = \exp\left[\frac{1}{N_0}\text{Re}\left\{d_{k,n}^*\theta_{k,n}^* y_{k,n} - \frac{1}{2}\lambda(k,k,n,n)|d_{k,n}|^2\right\}\right], \quad (16)$$

$$H_{k,n}^{a,b}(d_{k,n}, d_{k-a,n-b}) = \exp\left[-\frac{1}{N_0}\text{Re}\left\{\lambda(k,k-a,n,n-b) \cdot d_{k,n}^*\theta_{k,n}^* d_{k-a,n-b}\theta_{k-a,n-b}\right\}\right]. \quad (17)$$

After the factorization, the conditional PDF in Eq.(11) can be written as

$$p(\boldsymbol{r}|\boldsymbol{s}) \propto \prod_{k=1}^{M}\prod_{n=1}^{N}\left[F_{k,n}(d_{k,n})\prod_{a=1}^{L_1}H_{k,n}^{a,0}(d_{k,n}, d_{k-a,n}) \cdot \prod_{a=-L_1}^{L_1}\prod_{b=1}^{L_2}H_{k,n}^{a,b}(d_{k,n}, d_{k-a,n-b})\right]. \quad (18)$$

The final APP (A Posterior Probability) of the transmitted sequence is

$$p(\boldsymbol{s}|\boldsymbol{r}) \propto \prod_{k=1}^{M}\prod_{n=1}^{N}\left[P_{k,n}(d_{k,n})F_{k,n}(d_{k,n})\prod_{a=1}^{L_1}H_{k,n}^{a,0}(d_{k,n}, d_{k-a,n}) \cdot \prod_{a=-L_1}^{L_1}\prod_{b=1}^{L_2}H_{k,n}^{a,b}(d_{k,n}, d_{k-a,n-b})\right], \quad (19)$$

where $P_{k,n}(d_{k,n})$ is the a priori probability of the symbol $d_{k,n}$.

The FG corresponding to Eq.(19) can be depicted as shown in Fig.1 in the case where $L_1 = L_2 = 1$ and the number of interferers $L=8$. In the FG the 2D data $\{d_{k,n}\}$ are arranged in frequency priority order as [$d_{1,1}, d_{2,1}, \cdots, d_{M,1}, d_{1,2}, d_{2,2}, \cdots, d_{M,2}, \cdots$] such that they can be

regarded as a 1D data sequence, which helps to make the updating order clear if a serial updating schedule is used in the message passing.

3.3 Message passing

It can be seen in Fig.1 that the FG has cycles which means that the SPA cannot converge exactly, and furthermore, the APP will not be accurately calculated. However, a good approximation can be expected in practice because the length of the shortest cycle of the FG is six. The notations for the message passing, defined as in Ref.[6], are shown in Fig.2. Moreover, $V_{k,n}(d_{k,n})$ is defined as the product of all the incoming messages to the variable node $d_{k,n}$

$$V_{k,n}(d_{k,n}) = P_{k,n}(d_{k,n}) \, F_{k,n}(d_{k,n})$$
$$\cdot \prod_{a=1}^{L_1} \mu_{U,k,n}^{a,0}(d_{k,n}) \mu_{B,k,n}^{a,0}(d_{k,n})$$
$$\cdot \prod_{a=-L_1}^{L_1} \prod_{b=1}^{L_2} \mu_{U,k,n}^{a,b}(d_{k,n}) \mu_{B,k,n}^{a,b}(d_{k,n}) . \quad (20)$$

The message updating rules obey the following equations.

$$\mu_{F,k,n}^{a,b}(d_{k,n}) = \frac{V_{k,n}(d_{k,n})}{\mu_{B,k,n}^{a,b}(d_{k,n})} , \quad (21)$$

$$\mu_{D,k,n}^{a,b}(d_{k,n}) = \frac{V_{k,n}(d_{k,n})}{\mu_{U,k,n}^{a,b}(d_{k,n})} , \quad (22)$$

$$\mu_{U,k+a,n+b}^{a,b}(d_{k+a,n+b}) = \sum_{d_{k,n}} H_{k+a,n+b}^{a,b}(d_{k+a,n+b}, d_{k,n})$$
$$\cdot \mu_{F,k,n}^{a,b}(d_{k,n}) , \quad (23)$$

$$\mu_{B,k-a,n-b}^{a,b}(d_{k-a,n-b}) = \sum_{d_{k,n}} H_{k,n}^{a,b}(d_{k,n}, d_{k-a,n-b})$$
$$\cdot \mu_{D,k,n}^{a,b}(d_{k,n}) . \quad (24)$$

Both the serial and the parallel updating schedules presented in Ref.[6] can be applied in the message passing. The serial schedule results in a detection delay, but it converges faster than the parallel schedule.

In this paper, a serial schedule is considered because
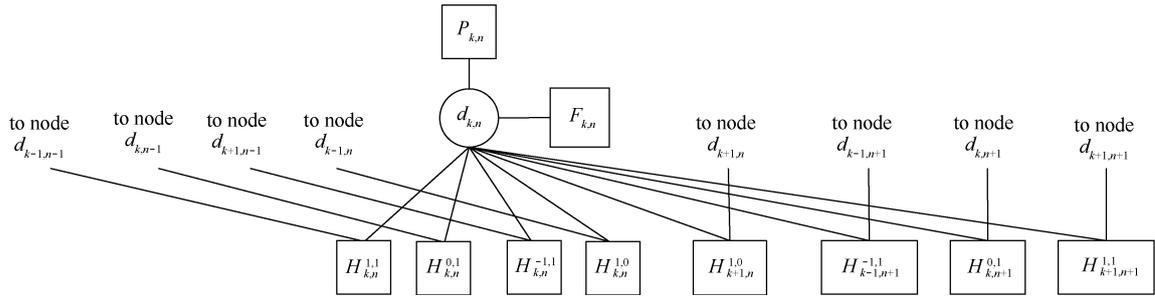
Figure 1 A section of the factor graph corresponding to Eq.(19) in the case where $L_1 = L_2 = 1$
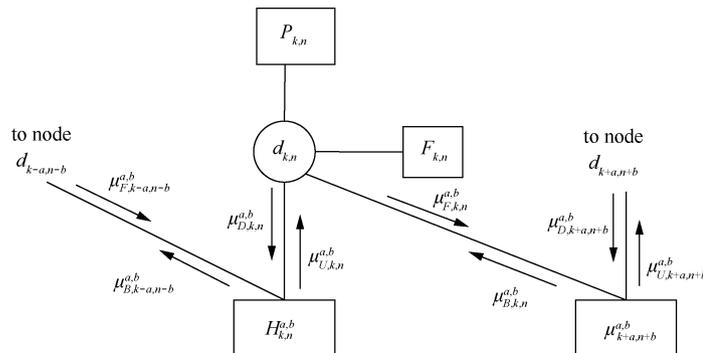
Figure 2 Part of the factor graph corresponding to Eq.(19) and the notations of the messages

of its better performance. The updating process consists of a forward recursion and a backward recursion. The forward recursion updates $\mu_{U,k,n}^{a,b}(d_{k,n})$, $V_{k,n}(d_{k,n})$ and $\mu_{F,k,n}^{a,b}(d_{k,n})$ successively for each $\{k, n\}$, using the priority order proposed above for k. The backward recursion updates $\mu_{B,k,n}^{a,b}(d_{k,n})$, $V_{k,n}(d_{k,n})$, and $\mu_{D,k,n}^{a,b}(d_{k,n})$ successively for each $\{k, n\}$ in the opposite order.

4 Extension to MIMO-FBMC

In MIMO systems, the conditional PDF $p(\boldsymbol{r}|\boldsymbol{s})$ is proportional to

$$
\exp\left(-\frac{\sum_{n_r}\sum_m \left|r_{n_r}(m)-\sum_{n_t}s_{n_r}(m)*c_{n_t,n_r}(m)\right|^2}{2N_0}\right)
$$

$$
= \exp\left[-\frac{1}{2N_0}\sum_{n_r}\left(\sum_m |r_{n_r}(m)|^2 + \sum_m\left|\sum_{n_t}s_{n_t}(m)*c_{n_t,n_r}(m)\right|^2\right.\right.
$$

$$
\left.\left. -2\mathrm{Re}\left\{\sum_m\sum_{n_t}r_{n_r}(m)(s_{n_r}(m)*c_{n_t,n_r}(m))^*\right\}\right)\right], \quad (25)
$$

where n_t and n_r are the index of the transmitting and receiving antennas, respectively, whereas N_t and N_r are defined as the total numbers of transmitting and receiving antennas, respectively.

Similar to the derivations in Section 3, we can express the second term in the argument of exponential Eq.(25) as

$$
\sum_m\left|\sum_{n_t}s_{n_t}(m)*c_{n_t,n_r}(m)\right|^2
$$

$$
= \sum_{n_t^1,n_t^2=1}^{N_t}\sum_{k_1,k_2=1}^{M}\sum_{n_1,n_2=1}^{N} d_{n_t^1,k_1,n_1}^*\theta_{k_1,n_1}^*
$$

$$
\cdot d_{n_t^2,k_2,n_2}\theta_{k_2,n_2}\lambda_{n_r}(n_t^1,n_t^2,k_1,k_2,n_1,n_2), \quad (26)
$$

with

$$
\lambda_{n_r}(n_t^1,n_t^2,k_1,k_2,n_1,n_2)
$$

$$
\approx \frac{1}{KM}\sum_{t=0}^{kM-1} H^*(l-k_1K)C_{n_1,n_t^1,n_r}^*(l)
$$

$$
\cdot H(l-k_2K)C_{n_2,n_t^2,n_r}(l)\mathrm{e}^{\mathrm{j}(n_1-n_2)\frac{M}{2}\frac{2\pi l}{KM}}, \quad (27)
$$

and the third term as

$$
2\mathrm{Re}\left\{\sum_m\sum_{n_t}r_{n_r}(m)(s_{n_t}(m)*c_{n_t,n_r}(m))^*\right\}
$$

$$
= 2\mathrm{Re}\left\{\sum_{n_t=1}^{N_t}\sum_{k=1}^{M}\sum_{n=1}^{N} d_{n_t,k,n}^*\theta_{k,n}^* y_{n_t,k,n,n_r}\right\}, \quad (28)
$$

with

$$
y_{n_t,k,n,n_r} = \frac{1}{KM}\sum_{l=0}^{KM-1} H^*(l-kK)C_{n,n_t,n_r}^*(l)R_{n,n_r}(l) \quad (29)
$$

After obtaining the representation of y and λ, we can factorize the APP by using the same method as in the single antenna case. The notations are shown in Fig.3. The main difference is that there is an additional dimension of interference caused by
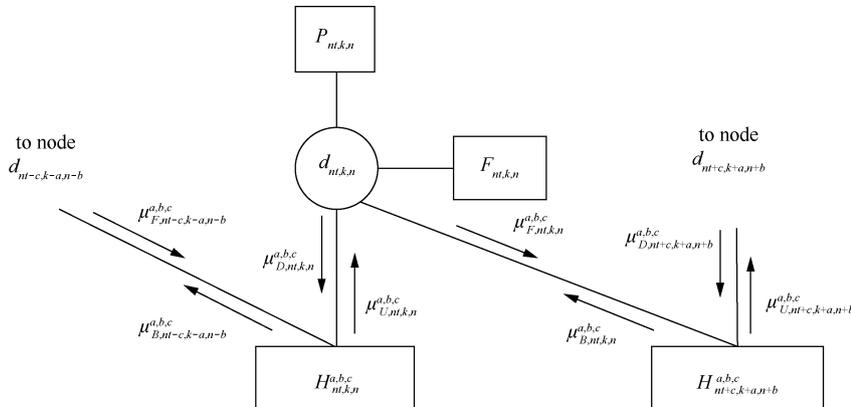

Figure 3 Part of the factor graph corresponding to Eq.(32) and the notations of the messages

the multiple transmitting antennas which increases the upper bound of the number of interferers L to $(2L_1+1)(2L_2+1)N_t-1$. Furthermore, the definition of the functions F and H should combine the multiple signals from the N_r receiving antennas, as follows

$$F_{n_t,k,n}(d_{n_t,k,n}) = \exp\left[\frac{1}{N_0}\sum_{n_r=1}^{N_r}\text{Re}\left\{d_{n_t,k,n}^*\theta_{k,n}^* y_{n_t,k,n,n_r}\right.\right.$$
$$\left.\left. -\frac{1}{2}\lambda_{n_r}(n_t,n_t,k,k,n,n)\,|\,d_{n_t,k,n}\,|^2\right\}\right],$$
(30)

$$H_{n_t,k,n}^{a,b,c}(d_{n_t,k,n},d_{n_t-c,k-a,n-b})$$
$$= \exp\left[-\frac{1}{N_0}\sum_{n_r=1}^{N_r}\text{Re}\left\{\lambda_{n_r}(n_t,n_t-c,k,k-a,n,n-b)\right.\right.$$
$$\left.\left. \cdot d_{n_t,k,n}^*\theta_{k,n}^* d_{n_t-c,k-a,n-b}\theta_{k-a,n-b}\right\}\right]$$
(31)

Finally, the APP of the transmitted sequence is

$$p(\boldsymbol{s}\,|\,\boldsymbol{r}) \propto \prod_{n_t=1}^{N_t}\prod_{k=1}^{M}\prod_{n=1}^{N}\left[P_{n_t,k,n}(d_{n_t,k,n})\,F_{n_t,k,n}(d_{n_t,k,n})\right.$$
$$\cdot \prod_{c=1}^{n_t-1}H_{n_t,k,n}^{0,0,c}(d_{n_t,k,n},d_{n_t-c,k,n})$$
$$\cdot \prod_{c=n_t-N_t}^{n_t-1}\left(\prod_{a=1}^{L_1}H_{n_t,k,n}^{a,0,c}(d_{n_t,k,n},d_{n_t-c,k-a,n})\right.$$
$$\left.\left. \cdot \prod_{a=-L_1}^{L_1}\prod_{b=1}^{L_2}H_{n_t,k,n}^{a,b,c}(d_{n_t,k,n},d_{n_t-c,k-a,n-b})\right)\right].$$
(32)

The FG corresponding to Eq.(32) can be depicted as in Fig.1. The 3D data $\{d_{n_t}, k, n\}$ are arranged in priority order of antenna, frequency and time as $[d_{1,1,1}, d_{2,1,1},\cdots, d_{N_t,1,1}, d_{1,2,1}, d_{2,2,1},\cdots, d_{N_t,2,1},\cdots]$ so that they can be regarded as 1D data. Then the message passing process is similar to that in the single antenna case.

5 Complexity analysis

5.1 FG-based detector

At the receiver, the baseband antenna signals are firstly processed by the AFB and subchannel filters to calculate $y_{k,n}$ in Eq.(14). When using the FS structure, a direct KM-point FFT operation whose complexity is $O(KM\log(KM))$ should be applied. Then the direct filtering will be operated in the frequency domain with a complexity of $O(W)$ for each sub-channel. As W is constantly proportional to K for a specific type of prototype filter, it follows that $O(W)=O(K)$. As the FFT is operated based on the symbol period, M symbols are processed together. Hence, the complexity of calculating $y_{k,n}$ is $O(K\log(KM))$ for each symbol. If the PPN structure is used, the complexity of the AFB will be $O(KM+M\log(M))$. As M W-tap time domain filters follow the AFB, the complexity of filtering is $O(KM)$ for one symbol period. Combining the AFB and sub-channel filters, we can write the complexity of calculating $y_{k,n}$ by using the PPN structure as $O(K+\log(M))$ for each symbol.

After the AFB and filters, we obtain the sequence $\{y_{k,n}\}$. However, to calculate the function nodes $F_{k,n}$ and $H_{k,n}^{a,b}$, we also need $\lambda(k,k-a,n,n-b)$. As discussed in Section 3.2, only W out of KM components in Eq.(15) are nonzero and need to be calculated. As W is proportional to K, the complexity of calculating $\lambda(k,k-a,n,n-b)$ is $O(K)$ for each pair of symbols. As all the other parameters are scalars, the complexity of calculating $F_{k,n}$ and $H_{k,n}^{a,b}$ is also $O(K)$ for each function node. Assuming that one symbol will be disturbed by L symbols on average, there will be L function nodes $H_{k,n}^{a,b}$ corresponding to one variable node $d_{k,n}$. As discussed in Section 3.1, the upper bound of L is limited by L_1 and L_2 so that $O(L)=O(L_1L_2)$. Thus the total complexity of computing the function nodes is $O(LK)=O(L_1L_2K)$ for one symbol. When the channel $c(m)$ is time-invariant, $\lambda(k,k-a,n,n-b)$ is also time-invariant and can be pre-computed. Then the complexity of calculating this term can be ignored when N is large and the total complexity of calculating the function nodes for each symbol will

be reduced to $O(L_1L_2)$.

The complexity of updating messages in the FG is proportional to the number of links between variable nodes and function nodes. For each variable node $d_{k,n}$, there are L function nodes $H_{k,n}^{a,b}$ connected to it. Hence, the number of passing messages for one variable node is proportional to L. With function nodes calculated as above, the updating process of each message is simply the sum and product of scalars whose complexity is $O(1)$. Therefore, the total complexity of updating messages for each symbol is $O(L_1L_2)$.

Combining all the steps above, we obtain the total complexity for each symbol as $O(K \log(KM) + L_1L_2K)$ for the FS structure and as $O(\log(M) + L_1L_2K)$ for the PPN structure. If the channel is time-invariant and $\lambda(k, k-a, n, n-b)$ is precalculated, the complexity can be reduced to $O(K \log(KM) + L_1L_2)$ for the FS structure and to $O(K + \log(M) + L_1L_2)$ for the PPN structure.

In contrast, for similar values of M and L_2, the length of the memory in the Viterbi or BCJR algorithm is $O(L_2M)$ which makes the complexity of the decoding process to be at least $O(2^{L_2M})$. The exponentially growing complexity is impractical and cannot be reduced even if most of the coefficients in the memory are zeros.

5.2 FG-based detector in MIMO systems

As the frequency width of the filter is the same as that in the single antenna case, the complexity of calculating the corresponding terms y_{n_t,k,n,n_r} and $\lambda_{n_r}(n_t, n_t-c, k, k-a, n, n-b)$ is unchanged. However, the use of N_r receiving antennas will result in N_r data streams and related calculations, which makes the total complexity N_r times larger for each symbol. Moreover, the number of interferers will increase to about N_t times because of the IAI between the data from different antennas.

Considering the above features of MIMO systems, the total complexity after combining all the steps of the receiver is $O(N_r K \log(KM) + N_r N_t L_1L_2K)$ for the FS structure and $O(N_r \log(M) + N_r N_t L_1L_2K)$ for the PPN structure. If the channel is time invariant and $\lambda_{n_r}(n_t, n_t-c, k, k-a, n, n-b)$ can be precalculated, the complexity can be reduced to $O(N_r K \log(KM) + N_r N_t L_1L_2)$ for the FS structure and to $O(N_r K + N_r \log(M) + N_r N_t L_1L_2)$ for the PPN structure.

It can be seen that the complexity of the proposed FG-based detector increases linearly with the number of transmitting antennas. In contrast, the memory length in the Viterbi or BCJR algorithm is $O(N_t L_2 M)$ and therefore, the complexity is at least $O((2N_r)^{N_t L_2M})$, which increases exponentially.

6 Simulation results

In this section, the performance of the FG-based detector is evaluated using computer simulations. Both PHYDYAS and IOTA pulses were considered as prototype filters in the simulation. For the IOTA filter, the design parameters from Ref. [14] were set as $\upsilon_0 = 1$, $\tau_0 = 1/2$, $\rho = 2$. The ITU Vehicular A multipath channel model is used[15], which comprised six paths with delays of [0, 0.31, 0.71, 1.09, 1.73, 2.51] μs and relative powers of [0, −1, −9, −10, −15, −20] dB. Assume that perfect channel state information was known. It is well known that when the number of carriers is large enough, the subbands will be approximately flat and the ISI will be small in the single antenna case which makes the performance of simple linear equalizers very good. Therefore, the superior performance of the FG-based MAP detector is more evident with fewer carriers in the single antenna case. The simulation parameters are listed as

- Overlapping factor: $K = 4$.
- Modulation scheme: OQAM.
- Sampling frequency: 15.36 MHz.
- Number of carriers: $M = 128$.
- Channel coding: LDPC code with $(n, k) = $ (8 176, 7 152).

With these parameters, the subbands were non-flat fading as shown in Fig.4. The results were obtained by Monte-Carlo simulations with independent trials. Here we show the performance of the FG-based SISO detector with K coefficients per subband. For the purpose of performance comparison, we also considered the MMSE and ZF equalizers.
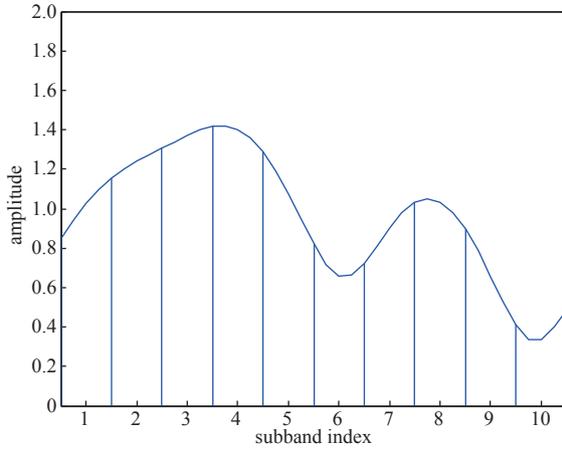

**Figure 4** Part of the non-flat fading subbands under the simulation conditions

For the factor graph, the serial message schedule was applied as proposed in Section 3.3. When the detector was provided with information from the decoder, only one iteration was executed inside it before outputting the extrinsic information. The number of turbo iterations between the detector and decoder was set as three. In addition, the messages in the FG of the detector were reset after the information exchange. The FG with cycles usually leads to the overestimation of the reliability of the propagated messages. Thus to improve the performance, we used a larger value of N_0 than the actual one. The N_0 was increased by 3 dB in the simulations. To mitigate the residual interference caused by some approximations, such as the circular convolution in calculating $\lambda(k_1, k_2, n_1, n_2)$, we set a threshold such that, even when E_b/N_0 >20 dB, N_0 could not be smaller than the variance of the noise when E_b/N_0 =20 dB.

6.1 Single antenna FBMC

Figs.5 and 6 report the BER performance of the FBMC system with IOTA and PHYDYAS prototype filters, respectively. In the figures we can see that the FG-based SISO detector combined with a turbo equalization structure had a better performance than the simple MMSE and ZF equalizers in coded systems. The performance of the ZF equalizer was not as good as those of MMSE and FG because of the residual interference caused by using circular
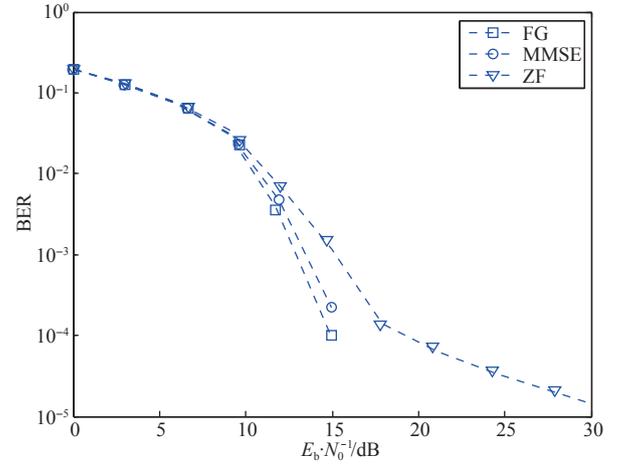
Figure 5 Coded BER performance of the FG-based detector and linear equalizers with IOTA prototype filter
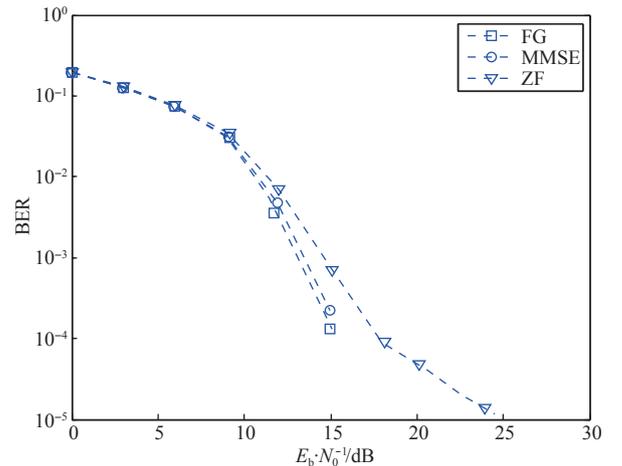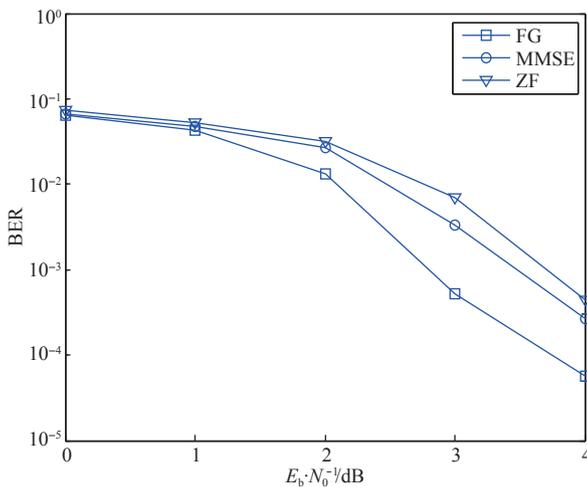

Figure 6 Coded BER performance of the FG-based detector and linear equalizers with PHYDYAS prototype filter

convolution to approximate the linear convolution of two pulses. The interference can be regarded as noise with constant variance whose effect will be predominant under high SNR (Signal-to-Noise Ratio). The MMSE equalizer offers a tradeoff between the noise and interference so that it achieves good performance at high SNR. However, the FG-based SISO detector can fully utilize the soft information and use the turbo equalization structure to achieve the best performance.

6.2 MIMO-FBMC

A 2×6 MIMO-FBMC system was considered in the simulation. The two antennas transmitted different data streams, which means that IAI would exist at the receiver. Figs.7 and 8 report the BER performance of the MIMO-FBMC system with IOTA and PHYDYAS prototype filters respectively. The three algorithms benefited from the spatial diversity of MIMO systems with the BER decreasing faster than that in the single antenna case. Similarly, the FG-based MIMO detector achieved better performance than the MMSE and ZF equalizers.

Figure 7　Coded BER performance of the FG-based detector and linear equalizers with IOTA prototype filter in MIMO systems
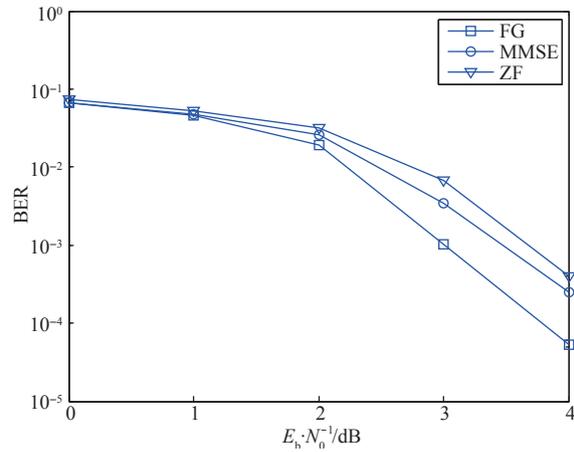
Figure 8　Coded BER performance of the FG-based detector and linear equalizers with PHYDYAS prototype filter in MIMO systems

7 Conclusion

In this paper, we considered the application of an FG-based SISO MAP detector to FBMC systems. The complexity of the proposed detector grows linearly with the number of interfering symbols, which makes it attractive for implementation under practical conditions. The detector can be naturally combined with a decoder to form a turbo structure to improve the performance. Moreover, we extended the proposed FG-based detector to MIMO-FBMC systems and the complexity increased linearly with the number of antennas. Simulation results showed that the FG-based detector can achieve better performance than the simple MMSE and ZF equalizers in coded systems.

References

[1]　FARHANG-BOROUJENY B. OFDM versus filter bank multicarrier[J]. IEEE signal processing magazine, 2011, 28(3): 92-112.

[2]　BELLANGER M, RENFORS M, IHALAINEN T, et al. OFDM and FBMC transmission techniques: A compatible high performance proposal for broadband power line communications[C]//International

Symposium on Power Line Communications and Its Applications (ISPLC), Rio de Janeiro, Brazil, 2010: 154-159.

[3] IHALAINEN T, IKHLEF A, LOUVEAUX J, et al. Channel equalization for multi-antenna FBMC/OQAM receivers[J]. IEEE transactions on vehicular technology, 2011, 60(5): 2070-2085.

[4] ZAKARIA R, LE RUYET D, BELLANGER M. Maximum likelihood detection in spatial multiplexing with FBMC[C]// European Wireless Conference (EW), Lucca, Italy, 2010: 1038-1041.

[5] KSCHISCHANG F R, FREY B J, LOELIGER H A. Factor graphs and the sum-product algorithm[J]. IEEE transactions on information theory, 2001, 47(2): 498-519.

[6] FERTONANI D, BARBIERI A, COLAVOLPE G. Novel graph-based algorithms for soft-output detection over dispersive channels[C]//Global Telecommunications Conference (GLOBECOM), New Orleans, USA, 2008: 1-5.

[7] UNGERBOECK G. Adaptive maximum likelihood receiver for carrier-modulated data-transmission systems[J]. IEEE transactions on communications, 1974, 22(5): 624-636.

[8] HASELMAYR W, ETZLINGER B, SPRINGER A. Factor-graph-based soft-input soft-output detection for frequency-selective MIMO channels[J]. IEEE communiations letters, 2012, 16(10): 1624-1627.

[9] DOUILLARD C, JEZEQUEL M, BERROU C, et al. Iterative correction of inter symbol interference: turbo-equalization[J].

European transactions on telecommunications, 1995, 6(5): 507-511.

[10] TUCHLER M, KOETTER R, SINGER A C. Turbo-equalization: principles and new results[J]. IEEE transactions on communications, 2002, 55(5): 754-767.

[11] BELLANGER M. FS-FBMC: an alternative scheme for filter bank based multicarrier transmission[C]//The 5th International Symposium on Communications, Control and Signal Processing (ISCCSP 2012), Rome, Italy, 2012: 1-4.

[12] MATTERA D, TANDA M, BELLANGER M. Frequency-spreading implementation of OFDM/OQAM systems[C]//International Symposium on Wireless Communication Systems (ISWCS 2012), Paris, France, 2012: 176-180.

[13] TANDA M, FUSCO T, RENFORS M, et al. Transmit/receive processing (single antenna), ICT-211887 PHYDYAS Deliverable 2.1[EB/OL]. http://www.ict-phydyas.org/delivrables/PHYDYAS-%20D2.1.pdf/view.

[14] SAHIN A, GUVENC I, ARSLAN H. A survey on multicarrier communications: prototype filters, lattice structures, and implementation aspects[J]. IEEE communications surveys & tutorials, 2014, 16(3): 1312-1338.

[15] MARQUES A G. Guidelines for evaluation of radio transmission technologies for IMT-2000: Recommendation ITU-R M.1225[S]. Minneapolis: Department of Electrical and Computer Engineering, University of Minnesota.1997.
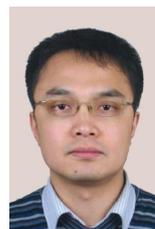
About the authors

CUI Fangyu received the B.S. degree in information science and electronic engineering from Zhejiang University, Hangzhou, China, in 2014. He is now a Ph.D. candidate of information science and communication engineering in Zhejiang University. His research interests include nonorthogonal modulation and detection. (Email: cfy531@zju.edu.cn)

ZHAO Minjian [corresponding author] received the M.Sc. and Ph.D. degrees in communication and information systems from Zhejiang University, Hangzhou, China, in 2000 and 2003, respectively. He is currently a professor with the Department of Information Science and Electronic Engineering, Zhejiang University, Zhejiang, China. His research interests include digital communication, software radio, and satellite communication. (Email: mjzhao@zju.edu.cn)

CAI Yunlong received the B.S. degree in computer science from Beijing Jiaotong University, Beijing, China, in 2004, the M.Sc.degree in electronic engineering from the University of Surrey, Guildford, U.K., in 2006, and the Ph.D. degree in electronic engineering from the University of York, York, U.K., in 2010. From February 2010 to January 2011, he was a postdoctoral fellow at the Electronics and Communications Laboratory of CNAM (the Conservatoire National des Arts et Metiers), Paris, France. Since February 2011, he has been with the Department of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China, where he is currently an associate professor. His research interests include spread spectrum communications, adaptive signal processing, multiuser detection, and multiple antenna systems. He has served on TPC (the Technical Program Committee) as member in various international conferences including for example IEEE GLOBECOM, IEEE ISWCS, and ITG WSA. He has also served as the Publicity Co-Chair for the fifth International Conference on WCSP (Wireless Communications and Signal Processing). (Email: ylcai@zju.edu.cn)

CHAMPAGNE Benoit received the B.Ing. degree in engineering physics from the École Polytechnique de Montréal, Montréal, QC, Canada, in 1983, the M.Sc. degree in physics from the Université de Montréal, Montréal QC, Canada, in 1985, and the Ph.D. degree in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 1990. From 1990 to 1999, he was an assistant and then associate professor at INRS Telecommunications, Université du Quebec, Montréal, QC, Canada. In 1999, he joined McGill University, Montreal, QC, Canada, where he is now a full professor within the Department of Electrical and Computer Engineering. He served as associate chairman of Graduate Studies in the Department from 2004 to 2007. His research focuses on the study of advanced algorithms for the processing of information bearing signals by digital means. His interests span many areas of statistical signal processing, including detection and estimation, sensor array processing, adaptive filtering, and applications thereof to broadband communications and audio processing, where he has co-authored more than 200 referred publications. His research has been funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada, the "Fonds de Recherche sur la Nature et les Technologies" from the Government of Quebec, as well as some major industrial sponsors, including Nortel Networks, Bell Canada, InterDigital, and Microsemi. He has been an Associate Editor for the IEEE Signal Processing Letters, the IEEE Transactions on Signal Processing and the EURASIP Journal on Applied Signal Processing. He has also served on the technical committees of several international conferences in the fields of communications and signal processing. (Email: benoit.champagne@mcgill.ca)