Channel Estimation for OFDM

Yinsheng Liu, Zhenhui Tan, Hongjie Hu, Leonard J. Cimini, Jr., *Fellow, IEEE*, and Geoffrey Ye Li, *Fellow, IEEE*

***Abstract*—*Orthogonal frequency division multiplexing* (OFDM) has been widely adopted in modern wireless communication systems due to its robustness against the frequency selectivity of wireless channels. For coherent detection, channel estimation is essential for receiver design. Channel estimation is also necessary for diversity combining or interference suppression where there are multiple receive antennas. In this paper, we will present a survey on channel estimation for OFDM. This survey will first review traditional channel estimation approaches based on *channel frequency response* (CFR). *Parametric model* (PM)-based channel estimation, which is particularly suitable for sparse channels, will be also investigated in this survey. Following the success of turbo codes and *low-density parity check* (LDPC) codes, iterative processing has been widely adopted in the design of receivers, and iterative channel estimation has received a lot of attention since that time. Iterative channel estimation will be emphasized in this survey as the emerging iterative receiver improves system performance significantly. The combination of *multiple-input multiple-output* (MIMO) and OFDM has been widely accepted in modern communication systems, and channel estimation in MIMO-OFDM systems will also be addressed in this survey. Open issues and future work are discussed at the end of this paper.**

***Index Terms*—OFDM, parametric model, iterative receiver, channel estimation, turbo principle, factor graph.**

I. Introduction

ORTHOGONAL frequency division multiplexing (OFDM) has been widely accepted in modern wireless communication systems due to its robustness against frequency selectivity in wireless channels [1]–[3]. In OFDM systems, a frequency selective channel is converted into a collection of flat fading channels, which can be compensated by simply using a one-tap equalizer. Hence, OFDM can greatly simplify equalizer design while enabling rather high data rates. As a result, OFDM has been standardized as a key physical-layer technique in many commercial systems [4]–[6].

As in many coherent communication systems, channel estimation is essential for receiver design in OFDM systems [7], [8]. However, channel estimation is a challenging problem in wireless systems due to the time variance and frequency selectivity of wireless channels. Although it is possible to avoid channel estimation by using differential detection, a 3–4 dB loss in *signal-to-noise ratio* (SNR) [7] will result.

There have been many works on channel estimation in OFDM systems [9], [10]. In general, channel estimation methods can be divided into four categories: traditional channel estimation based on the *channel frequency response* (CFR), *parametric model* (PM)-based channel estimation, *iterative channel estimation* (ICE), and channel estimation for *multiple-input multiple-output* (MIMO)-OFDM systems. CFR-based channel estimation can be considered as a basic approach. It gained attention in 1990's [11], and many commercial systems have adopted it due to its simplicity [4]. By utilizing a particular channel model or decided data symbols, the PM-based approaches or the ICE-based approaches can provide better performance. In addition, the principles of the PM-based approaches and the ICE-based approaches are quite different from those of the traditional CFR-based channel estimation approaches. Therefore, we regard the PM-based approaches and the ICE-based approaches as advanced channel estimation techniques.

Traditional channel estimation in OFDM systems is based on estimating the CFR [11], [12]. Typically, the CFRs of pilot symbols or training symbols are estimated first, then the CFRs of data symbols can be obtained through decision-based tracking or interpolation.

For PM-based approaches [13], the channel is modeled using a set of parameters. By estimating these parameters, the channel response can be reconstructed in the receiver. PM-based approaches are particularly suitable for sparse channels where the channel response can be modeled by a few dominant paths, typically two to six [14]. The dimension of the channel correlation matrix can be greatly reduced when it is constructed based on PM; correspondingly, the channel estimation performance can be improved.

Along with the success in error-control coding techniques, iterative processing has been extended to the overall design of the receiver. For example, the turbo receiver originated from the well-known turbo code first proposed in [15]. To the best of our knowledge, a turbo receiver is first used for OFDM in [16], where channel estimation works in an iterative manner. Another successful error-control code, proposed in [17]–[19], is the *low-density parity-check* (LDPC) code. The corresponding decoding algorithm is based on a factor graph representation of the LDPC code and the sum-product algorithm [20]–[22]. Similar to the turbo principle, a factor graph can also be used for the design of receivers [23]. Since the sum-product algorithm works iteratively, channel estimation in factor graph receivers also works in an iterative manner.

Manuscript received June 1, 2013; revised October 25, 2013 and January 13, 2014; accepted April 6, 2014. Date of publication May 12, 2014; date of current version November 18, 2014. This work was supported in part by the Fundamental Research Funds for the Central Universities under Grants 2013YJS022 and 2014YJS030 and in part by the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University under Grant RCS2010ZT011.

Y. Liu and Z. Tan are with Institute of Broadband Wireless Mobile Communications and State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China (e-mail: 09111035@bjtu.edu.cn; zhhtan@bjtu.edu.cn).

H. Hu is with Huawei Technologies (China) Company, Ltd., Shanghai 201206, China (e-mail: huhongjie@huawei.com).

L. J. Cimini, Jr., is with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716 USA (e-mail: cimini@udel.edu).

G. Y. Li is with the School of Electrical and Computer Engineering (ECE), Georgia Institute of Technology, Atlanta, GA 30332-0250 USA (e-mail: liye@ece.gatech.edu).

Digital Object Identifier 10.1109/COMST.2014.2320074

It has been proved that using MIMO significantly increases channel capacity in a flat fading channel [24]. Since OFDM can provide flat fading channels over all subcarriers, it is natural to combine MIMO and OFDM to support high quality data transmission [25]. Therefore, MIMO-OFDM has been widely used in many wireless systems [4]. However, approaches for channel estimation in *single-input single-output* (SISO) OFDM cannot be directly applied in MIMO-OFDM where signals from multiple transmit antennas arrive at the receive antenna simultaneously. Hence, channel estimation should be reconsidered in MIMO-OFDM systems [26].

In this survey, the topics above will be covered to provide a complete picture of channel estimation in OFDM systems. As far as we know, iterative channel estimation is not well surveyed in existing work. We will therefore give more attention to iterative channel estimation as this emerging receiver structure can significantly improve the system performance. The application of channel estimation approaches in current commercial system will be also discussed.

The rest of this paper is organized as follows. A description of system models is given in Section II. First, CFR-based channel estimation is presented in Section III, and then PM-based channel estimation is presented in Section IV. After that, iterative channel estimation is described in Section V. The three approaches above are also used in MIMO-OFDM systems, which is discussed in Section VI. Applications of channel estimation approaches are addressed in Section VII to bridge theory and practice. Future work and open issues are discussed in Section VIII. Finally, summaries are given in Section IX.

Abbreviations in this paper are listed in the following (alphabetical order):

AIC	Akaike information criterion
CDMA	code division multiple access
CFR	channel frequency response
CIR	channel impulse response
CP	cyclic prefix
DCT (IDCT)	(inverse) discrete cosine transform
DD-CE	decision-directed channel estimation
DFT (IDFT)	(inverse) discrete Fourier transform
EM	expectation maximization
ESPRIT	estimate signal parameters via rotational invariance technique
EXIT	extrinsic information transfer
FFT	fast Fourier transform
ICE	iterative channel estimation
ICI	inter-carrier interference
LDPC	low-density parity-check
LMMSE	linear minimum mean-square error
LS	least square
LTE	long-term evolution
MC	Monte Carlo
MDL	minimum description length
MIMO	multiple-input multiple-output
ML	maximum likelihood
MP	matching pursuit
MST	most significant tap
MUSIC	multiple signal classification
OFDM	orthogonal frequency division multiplexing
PA-CE	pilot-assisted channel estimation
PAPR	peak-to-average power ratio
PDA	probabilistic data association
PM	parametric model
RB	resource block
RE	resource element
SAGE	space alternating generalized expectation maximization
SISO	single-input single-output
SNR	signal-to-noise ratio
SP-CE	superimposed-pilot based channel estimation
STBC	space-time block coding
SVD	singular value decomposition
TTI	transmission time inverval
WSSUS	wide-sense stationary uncorrelated scattering
WiMAX	worldwide interoperability for microwave access
WLAN	wireless local area network

II. System Model

The time-varying *channel impulse response* (CIR) of a wireless channel can be modeled as [7]

$$h(t,\tau) = \sum_l \alpha_l(t) g(\tau - \tau_l), \tag{1}$$

where τ_l is the delay of the lth path, $\alpha_l(t)$ is the corresponding complex amplitude, and $g(\tau)$ is the impulse response of the transmit and receive filter. Assuming that every path has the same normalized correlation function, $r_t(\Delta t)$, then we have

$$r_{\alpha_l}(\Delta t) = \mathrm{E}\left\{\alpha_l(t+\Delta t)\alpha_l^*(t)\right\} = \sigma_l^2 r_t(\Delta t), \tag{2}$$

where σ_l^2 is the power of the lth path.

Given (1), the time-varying CFR at time t is

$$H(t,f) = G(f) \sum_l \alpha_l(t) e^{j2\pi f \tau_l}, \tag{3}$$

where $G(f)$ is the frequency response of the transmit and receive filters which can be assumed to be constant in the presence of guard subcarriers on both sides of the spectrum [27]. For *wide-sense stationary uncorrelated scattering* (WSSUS) channels [28], the correlation of the time-varying CFR can be represented as the product of the time domain correlation function and a frequency domain correlation [7], that is

$$\begin{aligned} r_H(\Delta t, \Delta f) &= \mathrm{E}\left\{H(t+\Delta t, f+\Delta f) H^*(t,f)\right\} \\ &= r_t(\Delta t) r_f(\Delta f) \end{aligned} \tag{4}$$

where $(\cdot)^*$ denotes the conjugate of complex, and

$$r_f(\Delta f) = \sum_l \sigma_l^2 e^{-j2\pi \Delta f \tau_l} \tag{5}$$

denotes the frequency domain correlation.

In an OFDM system, let T denote the symbol duration, then $1/T$ will be the subcarrier spacing, based on the orthogonal requirement. From (3), the CFR at the kth subcarrier of the nth OFDM symbol will be

$$H_{n,k} = H(nT, k/T). \tag{6}$$

Correspondingly, the correlation function of the channels for different symbols and subcarriers can be written as

$$r_H[n,k] = r_t[n]r_f[k] \tag{7}$$

where $r_t[n] = r_t(nT)$ and $r_f[k] = r_f(k/T)$.

Inter-carrier interference (ICI) due to time variations in the wireless channel is small for most practical vehicular speeds. As a result, the channel can be assumed as constant over one OFDM symbol, and thus ICI can be omitted for practical OFDM systems [29]. With this assumption, the received frequency domain signal, after removing the *cyclic-prefix* (CP) and applying the *inverse discrete Fourier transform* (IDFT), can be represented as

$$Y_{n,k} = H_{n,k}X_{n,k} + W_{n,k}, \tag{8}$$

where $W_{n,k}$ and $X_{n,k}$ denote, respectively, additive Gaussian noise with zero mean and variance σ_w^2, and a transmitted frequency domain signal at the kth subcarrier of the nth symbol.

Note that the frequency domain transmitted signal, $X_{n,k}$, can be either a modulated data symbol or a known pilot symbol. For *single-carrier frequency division multiple-access* (SC-FDMA) transmission [30], it can also be DFT-spread data signals. When SC-FDMA transmission is implemented using DFT-spread OFDM, as in LTE [4], it can be regarded as frequency domain transmission. The transmitted symbols are first spread using DFT, and then fed to the IDFT modulation. This is actually similar to OFDM transmission. As a result, channel estimation for SC-FDMA is similar to that for OFDM, and will not distinguish between these two transmission schemes in the remainder of this paper.

III. CFR-Based Channel Estimation

For CFR-based channel estimation, two basic estimation algorithms can be used, depending on whether statistical knowledge about the CFR is available or not.

Without statistical knowledge, CFRs can be treated as deterministic but unknown. In this case, *least-square* (LS) estimation can be used as it requires no statistical information about the CFRs. Following [31], LS estimation minimizes:

$$\hat{\mathbf{H}}_{\mathrm{LS}} = \arg\min_{\{\mathbf{H}\}} \|\mathbf{Y} - \mathbf{X}\mathbf{H}\|^2, \tag{9}$$

where $\mathbf{Y}$ and $\mathbf{H}$ are the received signal vector and the CFRs in vector form, represented as

$$\mathbf{Y} = [Y_0, Y_1, \ldots, Y_{K-1}]^{\mathrm{T}}$$
$$\mathbf{H} = [H_0, H_1, \ldots, H_{K-1}]^{\mathrm{T}},$$

with K the number of subcarriers. $\mathbf{X}$ is a $K \times K$ diagonal matrix with (k,k)th element given as

$$[\mathbf{X}]_{(k,k)} = X_k, \tag{10}$$

with X_k a known training or pilot symbol. Note that, in the description above, for ease of presentation, we only consider one single OFDM symbol. In the case of white Gaussian noise, it is easy to show that LS estimation is equivalent to *maximum-likelihood* (ML) estimation. For the classical estimation problem, ML estimation is the optimal approach and it can achieve the *Cramer-Rao bound* (CRB) [32]. Since LS (or ML) estimation requires no knowledge about the channel statistics, the estimation performance is generally not good enough. In addition, LS estimation may suffer from noise enhancement when the channel matrix has a large condition number.

For Bayesian estimation, the CFRs are considered as random variables with known statistics. The estimation performance can be greatly improved by exploiting this statistical information. In this case, a linear *minimum mean-square error* (LMMSE) estimator, which is designed to minimize $\mathrm{E}\{\|\widehat{\mathbf{H}} - \mathbf{H}\|^2\}$, is the optimal linear estimator. Direct calculation of this minimization yields the LMMSE estimates

$$\widehat{\mathbf{H}}_{\mathrm{LMMSE}} = \mathbf{R}_{\mathrm{H}}\left(\mathbf{R}_{\mathrm{H}} + \frac{1}{\gamma}\mathbf{I}\right)^{-1}\widehat{\mathbf{H}}_{\mathrm{LS}}, \tag{11}$$

where $\mathbf{R}_{\mathrm{H}} = \mathrm{E}(\mathbf{H}\mathbf{H}^{\mathrm{H}})$ is the channel correlation matrix, and $\gamma = \sigma_x^2/\sigma_w^2$ is the *signal-to-noise ratio* (SNR). Although the LMMSE estimator improves the performance significantly, the complexity is also increased due to the need to perform a matrix inversion.

Two basic strategies for OFDM channel estimation are *decision-directed channel estimation* (DD-CE) and *pilot-assisted channel estimation* (PA-CE). Typical pilot placements for both strategies are shown in Fig. 1. Besides DD-CE and PA-CE, channel estimation can also be based on superimposed pilots [33]–[38], which is called *superimposed-pilot based channel estimation* (SP-CE).

A. DD-CE

For DD-CE, the CFRs of the training symbols are first estimated, and then used to detect the subsequent data symbols. The detected symbol can be fed back to track the time-varying CFR.

LMMSE and LS estimation of CFRs are first proposed in [11]. To reduce the computational complexity, a modified LMMSE estimator is also studied where only the significant channel taps are taken into account. LMMSE estimation can also be derived through a low-rank estimator using a *singular-value-decomposition* (SVD) of the channel correlation matrix, $\mathbf{R}_{\mathrm{H}}$ [12]. In [12], only the frequency domain correlation is considered for channel estimation. Correlations in both the time domain and frequency domain are taken into account in [7]. In [7], [11], [12], and [39], the feedback symbols are obtained by hard decisions; feedback symbols can also be obtained by iteratively employing soft decisions and channel estimation, e.g., [40]–[42].

The main problem in DD-CE is error propagation, which becomes more of an issue with increasing Doppler frequency [9]. This problem can be solved by periodically inserting more training symbols in the transmitted frame [43]. In this case, DD-CE is actually equivalent to PA-CE with block-type pilots [44]. Robust statistics can also be used to alleviate the error propagation of DD-CE in fast fading channels [45].

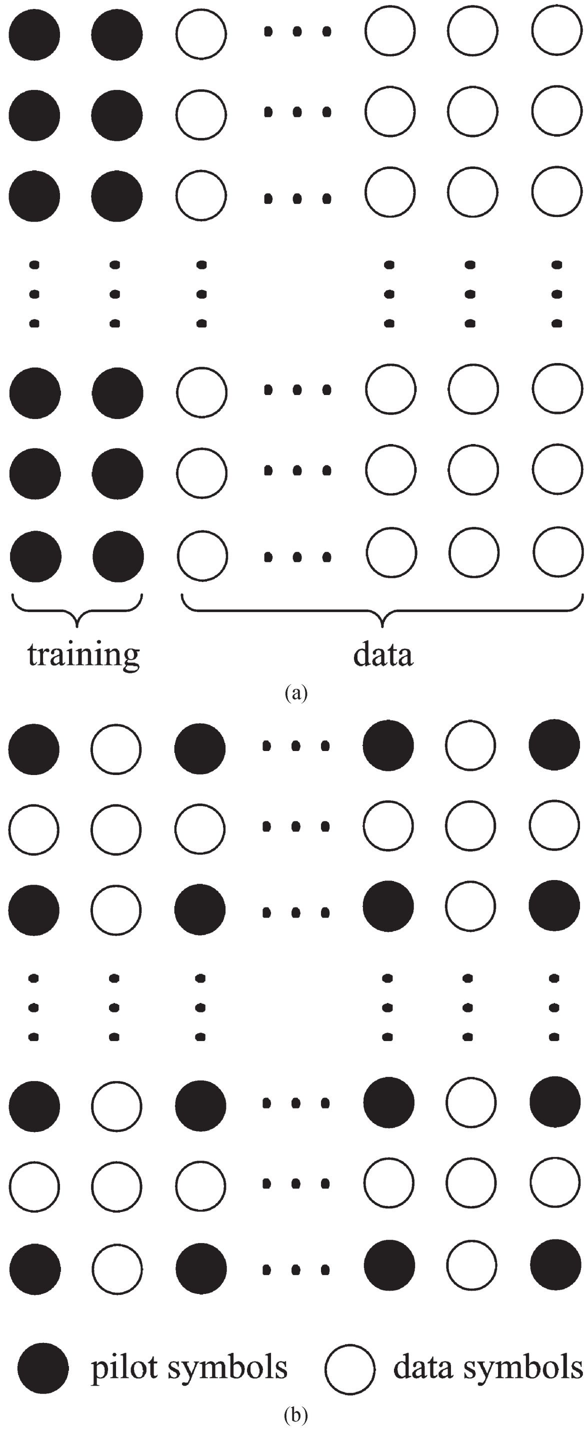


Fig. 1. Typical pilot symbol or training symbol placement. (a) DD-CE. (b) PA-CE.

B. PA-CE

For PA-CE, the CFRs of the pilot symbols are first estimated using LS estimation, then the CFRs corresponding to the data symbols can be obtained through interpolation. Inserting pilots in sequential OFDM symbols is equivalent to sampling the time-frequency channel, $H(t, f)$. Consequently, interpolation and pilot design are two main issues for PA-CE.

With the LS estimates of pilot symbols, the optimum interpolator is a *two-dimensional* (2-D) LMMSE interpolator [8], [46]. However, the complexity of 2-D LMMSE is too high and thus hard to implement in practical receivers. In addition, the statistical information required in the 2-D LMMSE interpolator is not always available in the receiver or it may change due to the time-variation of the wireless channel. Interpolators without statistical knowledge of the CFRs are studied in [3], [47]–[52]. Piecewise constant interpolation and piecewise linear interpolation, proposed in [3], [47], are the simplest interpolators. In highly frequency selective channels, however, more pilots are required for these interpolators to achieve satisfactory performance. The accuracy of interpolation can be further improved by using higher order interpolators, e.g., spline interpolation [48]–[50]. Given statistical knowledge of the channel, two *one-dimensional* (1-D) LMMSE interpolators are used in [47], [53] to reduce the complexity of the 2-D LMMSE interpolator. Two 1-D LMMSE interpolators include a time domain LMMSE interpolator and a frequency domain LMMSE interpolator. Using the 2×1 LMMSE interpolators, the complexity can be greatly reduced, with a small performance degradation. When channel knowledge is available, polynomial fitting can be also used for interpolation [51], [52].

There are two basic pilot patterns: block and the comb [44], [50], [54]. For the block-type pilot pattern, all subcarriers in specific time or frequency periodically transmit a block of pilot symbols. Channel responses for subcarriers between two pilot blocks can be obtained by interpolation or decision-directed estimation. Both LS and LMMSE can be used for channel estimation in the block-type pilot pattern [11], [12]. For a comb-type pilot pattern, the pilot symbols are spread among data subcarriers, at different subcarriers in different OFDM blocks. Usually, the pilot subcarriers are equally spaced in the frequency domain. Different interpolation techniques can be used to recover the channel responses at the data subcarriers [44], [55]. Different criteria have been adopted for optimizing the pilot pattern. It is found in [56] that equally-spaced pilot subcarriers can minimize the MSE of CIR estimation for a single OFDM symbol. In [57], it is shown that the placement of pilot subcarriers should also take into account the fourth-order moments of the statistics of channel responses at different times and frequencies. Pilot design to minimize the error probability or maximize the channel capacity in the presence of channel estimator error has been discussed in [58] and [59], respectively. Power allocation and pilot density have been investigated in [56]–[59].

C. SP-CE

Besides DD-CE and PA-CE, channel estimation can also be based on superimposed pilots where the pilot symbols are added onto the transmitted information symbols [33]–[38]. Since SP-CE does not need extra bandwidth for pilot symbols, the spectral efficiency can be improved accordingly. By including cyclostationary statistics at the transmitter through periodic pilot sequence, channel estimation can be conducted using the first-order statistical information at the receiver [36], [37]. If the information symbols are regarded as additive noise, an LS estimator can also be used to obtain the channel response, and the subsequent symbol detector works in a decision-feedback manner [60]. High PAPR is one of the main disadvantages in

OFDM systems. In [33], [38], it is found that the PAPR of the transmitted signal can be reduced by judiciously selecting the superimposed pilot sequence. It is also possible to combine the superimposed pilot-based channel estimation with the Viterbi algorithm [35].

D. Comparison

DD-CE is particularly suitable for burst transmission where the channel response is usually stationary or quasi-static. In training mode, the initial transmission is carrier out using training symbols. In data mode, the estimated channel can be tracked using decision-feedback channel estimation. Since the decision results are used as *known* training symbols, DD-CE can provide high spectral efficiency in quasi-static channels. However, when the channel varies rapidly with time, error propagation may lead to severe performance reduction [45].

When the channel is rapidly varying with time, the performance of channel estimation can be improved using PA-CE since the pilots are distributed more uniformly in the time-frequency plane. In traditional PA-CE, however, only the pilots are used for channel estimation, and many pilots are required to achieve satisfactory performance. PA-CE thus reduces the spectral efficiency compared with DD-CE where decision results can be treated as training symbols. This problem can be solved by iterative channel estimation where the decision results in PA-CE are used to re-estimate channel. We will cover this topic in Section V.

The major advantage of SP-CE is that it causes no loss in spectral efficiency; so this approach is especially suitable for bandwidth-limited systems. However, it is not suitable for power limited systems because part of the power is allocated to the pilots [34]. In other words, using an explicit pilot or a superimposed pilot depends on whether the system is power-limited or bandwidth-limited.

E. Transform Domain Techniques

Usually, the number of significant CIR taps is much smaller than the number of subcarriers [11]. This property has been widely exploited through transform domain techniques in both DD-CE and PA-CE. In [11], by transforming the CFR to a CIR, the computational complexity of the LMMSE estimator can be reduced if only the significant CIR taps are considered. A similar effect can be achieved by transforming the CFR into the eigenvalue domain using SVD [12]. Besides reducing complexity, transform domain techniques can also be used for noise reduction. In [61], DFT-based noise reduction is used to improve the accuracy of channel estimation; however, discontinuities at the edges of the spectrum may lead to the Gibbs phenomena and thus degrade the performance. Hence, other works resort to the *discrete cosine transform* (DCT) for noise reduction because it can maintain continuity at the edges of the spectrum [62]–[64]. For PA-CE, by inserting zeros at the CFRs corresponding to the data symbols, interpolation can be carried out using DFT-based transform domain techniques [65].

CFR-based channel estimation has gained a lot of attentions in the 1990's [11]. In OFDM systems, the modulation symbols are transmitted and received in the frequency domain. It is therefore natural to estimate the CFRs for coherent detection. In addition, through low complexity filtering [12], [47], [53], the complexity for the CFR-based approach can be greatly reduced. As a result, CFR-based channel estimation had become a popular technique and has been widely used in commercial systems [4]–[6]. However, it is possible to improve the channel estimation performance by using more advanced channel estimation techniques. As a result, CFR-based channel estimation can be considered as an elementary approach.

IV. PM-Based Channel Estimation

In PM-based channel estimation, the purpose is to estimate the parameters which characterize the channel response, e.g., the number of paths, path delays, and path gains in (3). The dimension of the channel correlation matrix can be greatly reduced if it is constructed based on PM, leading to significantly improved channel estimation performance when the PM of channel is valid.

To determine the number of paths, different criteria for model order selection are available in existing literature [31]. For path delay estimation, estimating the path delays using frequency domain pilots is equivalent to estimating the arrival angle using an antenna array [66]. Thus, well-known signal processing techniques, e.g., *estimation of signal parameters via rotational invariance techniques* (ESPRIT) [67], can be adopted for this purpose. With the estimates of path delays, the path gains can be obtained using typical linear estimators.

The key task in PM-based approaches is to estimate path delays using the ESPRIT algorithm. As described in Section II, the CFR at the kth subcarrier of the nth OFDM symbol is

$$H[n,k] = \sum_{l} \alpha_l(nT) e^{-j\frac{2\pi\tau_l k}{T}}, \tag{12}$$

which can be written in vector form as

$$\mathbf{H}[n] = \begin{bmatrix} H[n,0] \\ H[n,1] \\ \vdots \\ H[n,K-1] \end{bmatrix}. \tag{13}$$

Thus, frequency domain channel correlation can be also given in matrix form as

$$\mathbf{R}_f = \mathrm{E}\left\{\mathbf{H}[n]\mathbf{H}^{\mathrm{H}}[n]\right\} = \mathbf{V}\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}}, \tag{14}$$

where $\mathbf{\Sigma}$ is a diagonal matrix with $[\mathbf{\Sigma}]_{(l,l)} = \sigma_l^2$, and $\mathbf{V} = [\mathbf{v}(\tau_1), \mathbf{v}(\tau_2), \ldots, \mathbf{v}(\tau_L)]$ with $\mathbf{v}(\tau_l) = [1, e^{-j(2\pi\tau_l/T)}, \ldots, e^{-j(2\pi\tau_l(K-1)/T)}]^{\mathrm{T}}$ the steering vector. Note that $\mathbf{R}_f$ is independent of the time index, n, for WSSUS channels. In general, the correlation matrix, $\mathbf{R}_f$, is unknown to the receiver, but can be estimated through time averaging

$$\mathbf{R}_f \approx \frac{1}{N}\sum_{n=0}^{N-1} \hat{\mathbf{H}}[n]\hat{\mathbf{H}}^{\mathrm{H}}[n] \tag{15}$$

where $\hat{\mathbf{H}}[n]$ denotes the LS estimate of the CFRs. In this case, $\mathbf{R}_f$ can be factored, through eigenvalue decomposition, as

$$\mathbf{R}_f = [\mathbf{U}_{\mathrm{s}}, \mathbf{U}_{\mathrm{w}}] \begin{bmatrix} \mathbf{\Lambda}_{\mathrm{s}} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_{\mathrm{w}} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{\mathrm{s}}^{\mathrm{H}} \\ \mathbf{U}_{\mathrm{w}}^{\mathrm{H}} \end{bmatrix}, \tag{16}$$

where $\mathbf{U}_\mathrm{s}$ and $\mathbf{\Lambda}_\mathrm{s}$ are the eigenvector matrix and the eigenvalue matrix corresponding to the signal subspace, and $\mathbf{U}_\mathrm{w}$ and $\mathbf{\Lambda}_\mathrm{w}$ correspond to the noise subspace. The noise subspace is present because of the noise introduced by LS estimation.

If we define $\mathbf{V}_1$ and $\mathbf{V}_2$ as $(K-1) \times L$ matrices constructed by the first $K-1$ rows and the last $K-1$ rows of $\mathbf{V}$, respectively, we find that

$$\mathbf{V}_2 = \mathbf{V}_1 \mathbf{\Phi} \tag{17}$$

where $\mathbf{\Phi}$ is a diagonal rotation matrix with $(\mathbf{\Phi})_{(l,l)} = e^{-j(2\pi\tau_l/T)}$. The estimates of the path delays can be derived by obtaining the elements of $\mathbf{\Phi}$. For this purpose, we note, from (14) and (16), that the columns of $\mathbf{V}$ and the columns of $\mathbf{U_s}$ span the same signal space. Hence, there is an invertable linear transform, denoted as $\mathbf{T}$, that can map $\mathbf{U}_\mathrm{s}$ to $\mathbf{V}$. If we define $\mathbf{U}_1$ and $\mathbf{U}_2$ to be $(K-1) \times L$ matrices constructed by the first $K-1$ rows and the last $K-1$ rows of $\mathbf{U}_\mathrm{s}$, respectively, we obtain

$$\mathbf{V}_1 = \mathbf{U}_1 \mathbf{Q}, \text{ and } \mathbf{V}_2 = \mathbf{U}_2 \mathbf{Q}. \tag{18}$$

From (17) and (18), we find

$$\mathbf{U}_2 = \mathbf{U}_1 \mathbf{\Psi} \tag{19}$$

where $\mathbf{\Psi} = \mathbf{Q}\mathbf{\Phi}\mathbf{Q}^{-1}$. Hence, the elements of $\mathbf{\Phi}$ are exactly the eigenvalues of $\mathbf{\Psi}$, which can be estimated, from (19), as

$$\hat{\mathbf{\Psi}} = \left(\mathbf{U}_1^\mathrm{H} \mathbf{U}_1\right)^{-1} \mathbf{U}_1^\mathrm{H} \mathbf{U}_2. \tag{20}$$

Finally, we obtain

$$\hat{\tau}_l = -\frac{T}{2\pi} \angle \left\{ (\mathbf{\Phi})_{(l,l)} \right\} \tag{21}$$

where $\angle(\cdot)$ denotes the phase of the complex argument.

In essence, the ESPRIT algorithm utilizes the signal subspace for path delay estimation. Another well-known subspace-based approach is the *multiple signal classifier* (MUSIC) algorithm [68], which can also be used for path delay estimation. However, compared to the ESPRIT algorithm, the MUSIC algorithm requires an exhaustive search to obtain the estimates of the path delays, resulting in a significant increase in complexity. Therefore, most existing works use the ESPRIT algorithm for path delay estimation.

Above, we assume that all the frequency domain symbols are known. ESPRIT algorithm is first used in [13] for path delay estimation where pilots are multiplexed with unknown data symbols. This corresponds to a sampled CFR with a low sampling rate in the frequency domain. The ESPRIT algorithm can be still used as long as the number of pilots is larger than the number of significant paths. By averaging over many OFDM symbols, the channel correlation matrix can be estimated by obtaining LS estimates on the equispaced pilot subcarriers. The forward-backward approach [69] is used in [13] to improve the accuracy of the estimated correlation matrix. With the estimated path delays, the path gains are easy to obtain using the LMMSE algorithm. A delay locked loop is also employed to track the time variations of those delays.

Parameter estimation in uplink transmission is considered in [70]. Because the pilot allocation in the uplink transmission has irregular spacing, the averaging in [13] cannot be used. Instead, subspace-tracking [71] is used to obtain the required correlation matrix. In [72], by simulating the effect of the fading channel with hopping pilots, the convergence rate of the ESPRIT algorithm can be accelerated. In [73], two-stage ESPRIT algorithms are used to estimate the path delays and Doppler frequencies. To estimate the length of the CIR, different criteria are employed in the existing work. A *minimum description length* (MDL) criterion [74] is used in [13], [73], and the *Akaike information criterion* (AIC) [75] is used in [70]. In [72], the CIR length is simply chosen as the number of significant eigenvalues. Instead of estimating the path delays, variants of the PM-based approaches can also be found in [27], [76]–[78]. A sample spaced sparse channel is considered in [76]–[79]. For a sample spaced CIR, the paths delays are exactly integer multiple of the sample time. Therefore, estimating the path delays is equivalent to determining the path position in the range $[0, 1, \ldots, L-1]$ where L denotes the number of sample spaced paths. In [76], [77], the path positions can be obtained by iteratively using the AIC criterion. Alternatively, matching pursuit can be also employed for the same purpose [78], [79]. In [27], by noting that the delay-subspace is rank-deficient, subspace tracking is adopted to obtain the eigen-subspace of that delay-subspace instead of estimating the real path delays.

In general, the PM-based approach can improve the channel estimation performance when the number of unknown parameters is less than the dimension of the channel correlation matrix. This is usually the case for wireless channels in rural areas. To estimate the unknown parameters directly, subspace-based signal processing techniques have to be used, such as MUSIC and ESPRIT [67], [68]. Also, different criteria could be adopted to determine the number of parameters. Although there is significant performance improvement compared to the CFR-based approach [13], the disadvantages of the PM-based approach are also obvious. First, performance improvement can be obtained when the CIR is sparse in the time domain. This is usually the case in rural areas where the number of arrival paths is very small. However, most communications nowadays happen in urban areas. This restricts the use of the PM-based approach. Second, in the PM-based approach, the correlation matrix required by the subspace-based signal processing techniques is obtained by averaging over a long symbol sequence; this makes the PM-based approach only suitable for continuous transmission, such as digital broadcasting, and hardly used in burst communication systems, such as *Wireless Local Area Networks* (WLAN) [6].

V. Iterative Channel Estimation

The motivation for iterative channel estimation is to improve the accuracy of channel estimation using soft information about the data symbols. Although this can be achieved using the PM-based approaches described above, the requirement of a long symbol sequence for averaging or subspace tracking makes PM-based approaches only suitable for continuous transmission, e.g., digital broadcasting. Hence, iterative channel estimation is more suitable for cellular systems.

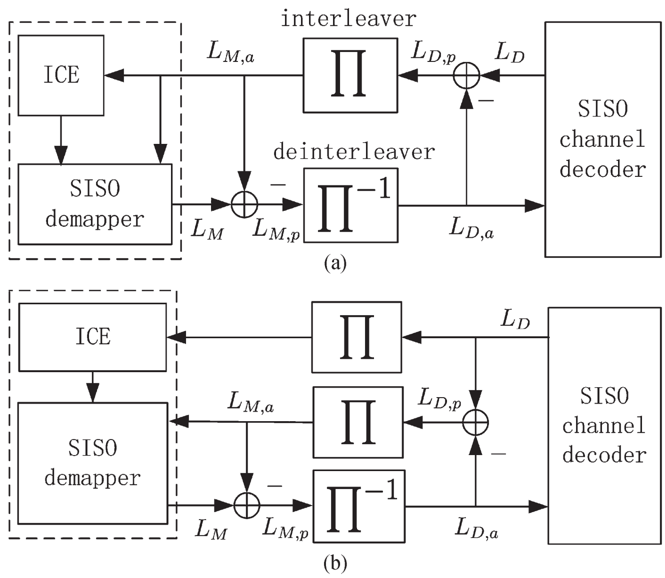


Fig. 2. (a) Typical turbo receiver where the extrinsic information is exchanged between the soft demapper and the soft decoder. (b) Turbo receiver with two kinds of feedback information. One is for channel estimation and the other is for the soft demapper.

The soft information required for CFR estimation can be obtained through decoder feedback or message-passing, depending on the receiver structure.

A. Iterative Channel Estimation for Turbo Receiver

A typical turbo receiver is shown in Fig. 2 [80]. The soft demapper in a turbo receiver calculates the log-likelihood ratios (soft information) on the modulation bits for each incoming symbol. The soft information is composed of an *a priori* part and an *a posteriori* part. Taking a *quadrature-phase-shift-keying* (QPSK) demapper, for example, the soft information for modulation bit $c_{n,k}^{(i)}$, where $i = 0, 1$ denotes the ith modulation bit for the QPSK symbol, is given as

$$L_M\left[c_{n,k}^{(i)}\right] = \ln \frac{\mathrm{P}\left\{c_{n,k}^{(i)} = 1 | Y_{n,k}, H_{n,k}\right\}}{\mathrm{P}\left\{c_{n,k}^{(i)} = 0 | Y_{n,k}, H_{n,k}\right\}} = L_{M,a}\left[c_{n,k}^{(i)}\right] + L_{M,p}\left[c_{n,k}^{(i)}\right], \tag{22}$$

with

$$L_{M,a}\left[c_{n,k}^{(i)}\right] = \ln \frac{\mathrm{P}\left\{c_{n,k}^{(i)} = 1\right\}}{\mathrm{P}\left\{c_{n,k}^{(i)} = 0\right\}}, \tag{23}$$

representing the *a priori* information provided by the decoder feedback, and $L_{M,p}[c_{n,k}^{(i)}]$ representing the *a posteriori* information provided by the soft demapper. After subtracting $L_{M,a}$ and deinterleaving, the output feeds the soft decoder as *a priori* information, as shown in Fig. 2(a).

In (22), perfect CFR is assumed. Otherwise, it has to be estimated using iterative channel estimation. Different from traditional channel estimation that makes use of only pilots or training symbols, iterative channel estimation has *a priori* information of the data symbols provided by the decoder feedback. Most of the work in this area has focused on how to use this information. We will review these works in terms of the following aspects.

1) Priori Information for Channel Estimation: Since there is some *a priori* information about data symbols from the decoder feedback, it is natural to determine first what the fed back information is. In [16], [81]–[84], only the *a posteriori* part of the soft information [cf. (22)] is fed back. This is exactly the situation shown in Fig. 2(a), where $L_{D,a}$ is canceled from the feedback path. In our view, those works can be more or less considered as extensions of turbo coding [15] or applications of the more general "turbo principle" [85], [86]. Following the "turbo principle," only extrinsic information is allowed to be exchanged between component devices while the *a priori* part is not included in the exchange loop. It is then a natural choice to feed back only the *a posteriori* part.

In other works, e.g., [40], [87]–[91], the *a priori* part and the *a posteriori* part are both fed back. In this case, the soft information is fed back for only channel estimation, while the demapper works without any *a priori* information. To also provide soft information for the demapper, different types of soft information can be fed back separately, as done in [92]. In [92], global soft information (including *a priori* and *a posteriori*) and the *a posteriori*-only part are both fed back in separate paths, as shown in Fig. 2(b). The global soft information is fed back for channel estimation, while the *a posteriori* part is fed back for the soft demapper. In [93] and [94], feedback is generated through hard decisions and re-encoding; although this idea is straightforward, no soft information is fed back.

2) Soft and Hard Mapping: The key for channel estimation in a turbo receiver is how to use the *a priori* information; generally, *a priori* information can be used to yield either soft symbols [16], [81], [95], [96], or hard symbols [80], [88], [89], [97].

A soft symbol is the expectation of the transmitted symbol. Taking QPSK modulation, for example, the regenerated soft symbol can be represented as

$$\tilde{X}_{n,k} = \sum_{i,j} P_{ij} Z_{ij} \quad i, j = 0, 1, \tag{24}$$

where $P_{ij} = P(c^{(0)} = i, c^{(1)} = j)$ is the *a priori* probability obtained from the decoder feedback and Z_{ij} is the constellation point corresponding to $c^{(0)} = i$ and $c^{(1)} = j$. Note that when there is no *a priori* information, i.e., $P_{ij} = 0.25$ for all i, j, the regenerated soft symbol is zero. The *a priori* information provides a regenerated soft symbol that is different from zero, and thus can be used for channel estimation.

Using the regenerated soft symbol, the channel estimate at the (n, k)th data subcarrier is given as (e.g., [16], [41], [95])

$$\tilde{H}_{n,k} = \frac{Y_{n,k}}{\tilde{X}_{n,k}}. \tag{25}$$

A 2-D symbol-wise Wiener filter is used in [16] to smooth the estimates, $\tilde{H}_{n,k}$, under the assumption that data symbols are recovered perfectly. An optimized smoothing filter is proposed in [80] to address estimation in the presence of symbol errors.

To reduce complexity, a sub-constellation containing less points (e.g., 4) is considered in [83] for higher order *quadrature-amplitude-modulation* (QAM), e.g., 64-QAM.

If the regenerated data symbol is directly applied in (25), channel estimation performance will degrade due to imperfect data symbol recovery [87] or noise enhancement [98]. The problem of biased estimation can be overcome by normalizing the average power of the regenerated soft symbol to unity [87]. In [81], instead of constructing the mean of the data symbol, the mean of the estimated CFRs is adopted, that is

$$\tilde{H}_{n,k} = \sum_{i,j} P_{ij} \frac{Y_{n,k}}{Z_{ij}}. \tag{26}$$

The estimate in (26) can avoid the noise enhancement since any point, Z_{ij}, in the constellation has a significant amplitude regardless of what the *a priori* probabilities are.

A hard symbol, on the other hand, is generated by choosing the element from the constellation that has the highest probability [80], [82], [89], [90], [97], that is

$$\hat{X}_{n,k} = \arg\max_{\{Z_{ij}\}} \{P_{ij}\}. \tag{27}$$

This ensures that the recovered symbol, $\hat{X}_{n,k}$, has significant amplitude for M-QAM or constant amplitude for M-ary *phase-shift-keying* (MPSK) and, thus, the noise enhancement can be avoided.

The major drawback of this hard remapping is that an error in the decision might propagate from one iteration to the next when the reliability of the decoder output is low. This problem can be addressed by applying a threshold test [97]. For this purpose, a cost function is specified that allows an assessment of whether regenerated symbols or pilot-assisted channel estimation should be used for the ith iteration, that is

$$\hat{H}_{n,k} = \begin{cases} Y_{n,k}/\hat{X}_{n,k} & \epsilon(Y_{n,k}/\hat{X}_{n,k}) \leq \epsilon\left(\hat{H}_{n,k}^{(0)}\right) \\ \hat{H}_{n,k}^{(0)} & \epsilon(Y_{n,k}/\hat{X}_{n,k}) > \epsilon\left(\hat{H}_{n,k}^{(0)}\right), \end{cases} \tag{28}$$

where $\hat{H}_{n,k}^{(0)}$ is the initial channel estimate, i.e., the estimate obtained using pilot symbols, and $\epsilon(Y_{n,k}/\hat{X}_{n,k})$ is the desired feedback quality indicator, which can be calculated using the soft information of the data symbols. Similar techniques are also adopted in [88] and [96]. In [88], erasure symbols are used when the log-likelihood-ratio is less than some predetermined threshold. The erasure symbols can be handled by replacing them with the nearest pilot symbol. In [96], zeros are inserted at the corresponding symbol positions where the decoded symbols do not have a sufficient level of confidence.

3) Multiple Iteration Oriented Approach: In the iterative approach, there is a temporary estimation at each iteration. It is therefore natural to consider using estimates from other iterations if the current estimate is not satisfactory [93], [94], [98].

Hard frequency replacement and soft frequency replacement proposed in [98] exploit the channel estimates from other iterations. Hard frequency replacement is a low complexity algorithm, originally proposed for mitigating the noise enhancement in (25). By using a threshold, λ, the channel estimates for the current iteration are obtained using

$$\tilde{H}_{n,k}^{(i)} = \begin{cases} Y_{n,k}/\tilde{X}_{n,k}^{(i)} & |\tilde{X}_{n,k}| \geq \lambda \\ \tilde{H}_{n,k}^{(i-1)} & |\tilde{X}_{n,k}| < \lambda. \end{cases} \tag{29}$$

In (29), determining the threshold is critical. In [98], the threshold is optimally determined using the analytically derived MSE. A technique similar to hard frequency replacement is described in [93], where the threshold is determined by simulation. Soft frequency replacement is another technique proposed in [98]. Instead of choosing between the two terms in (29), soft frequency replacement uses a weighted sum of both, with a weighting coefficient β

$$\tilde{H}_{n,k}^{(i)} = \beta Y_{n,k}/\tilde{X}_{n,k}^{(i)} + (1-\beta)\tilde{H}_{n,k}^{(i-1)}. \tag{30}$$

Estimates from other iterations are also exploited in [81], computing

$$\hat{H}_{n,k} = \max_{P(Z_{ij})} \left[P(Z_{ij})\frac{Y_{n,k}}{Z_{ij}}\right] + \left[1 - P(Z_{ij})\right] \hat{H}_{n,k}^{(0)}. \tag{31}$$

In this approach, the influence of the soft information from the error-correction decoder gradually increases and the influence of the first-round pilot-aided channel estimate decreases as the soft information gets better from iteration to iteration. A very similar method can be found in [92].

4) Expectation-Maximization (EM) Algorithm: In other work, e.g., [99]–[101], the EM algorithm is used for iterative channel estimation. The EM algorithm can perform ML estimation when there is incomplete data, Y (observable), and hidden data, X (unobservable) [102]. Generally, it consists of two steps: expectation and maximization. In the expectation step, the expectation function conditioned on the observed data, Y, and the previous estimate at the last iteration, $\boldsymbol{\theta}^{(i)}$, is found by

$$Q\left(\theta|\theta^{(i)}\right) = E_{X|y,\theta^{(i)}}\left[\log f(X, y|\theta)\right].$$

In the maximization step, the parameter is re-estimated by

$$\hat{\theta}^{(i+1)} = \arg\max_{\{\theta\}} Q\left(\theta|\theta^{(i)}\right).$$

In a turbo receiver, the hidden data X are the unknown data symbols, $X_{n,k}$, and θ is the unknown channel response, $H_{n,k}$. Note that the expectation step requires an average over all possible symbols using the *a priori* probabilities; these probabilities are exactly what the decoder feeds back in a turbo receiver. Hence, the EM algorithm is a good match for channel estimation in a turbo receiver [100].

It is generally difficult to analyze the convergence behavior of the turbo receiver. Therefore, many works use the *extrinsic information transfer* (EXIT) chart to determine the convergence. The EXIT chart is introduced in [103] and [104] to design signal constellations of an iterative demapping and decoding scheme. Then, it is used to analyze the convergence behavior of turbo codes in [105]. For a turbo receiver with ICE, the EXIT chart is used to analyze the convergence of iterative decoding and demapping in [16], [81], and [92].

In essence, the EM algorithm approaches ML estimation in an iterative manner. Alternatively, ML estimation can also be achieved using the *Monte Carlo* (MC) algorithm [106], [107], which utilizes sequences of random signals to approximate the expected estimator output. When the *a priori* distributions are unknown, the MC algorithm has been used in [108]–[110] for iterative channel estimation and symbol recovery.

B. Iterative Channel Estimation in Factor Graph Receiver

In a turbo receiver, the reliability information about decided symbols is fed back to improve the channel estimation performance. Actually, it is also possible to provide reliability information about channel estimation. This is exactly what the factor graph receiver does. Accordingly, channel estimation in a factor graph receiver is also called soft channel estimation, as opposite to the channel estimation in a turbo receiver which provides only an estimation result without reliability information [111]–[113]. In this section, we review iterative channel estimation in a factor graph receiver. First, we introduce the factor graph.

1) Factor Graph Receiver: Factor graph and corresponding sum-product algorithm are able to describe a wide variety of well-known algorithms [22]; the connection between factor graph and the more general Bayesian network [114] is also obtained in [22]. Factor graph is a visual representation of the sum-product algorithm [22], [115], which is designed to handle global functions of multiple variables. It can be considered as an equivalence to the Bayesian network [114], but due to its simplicity, factor graph has gain more attention in the design of many algorithms in artificial intelligence, signal processing, and digital communications [115]. Although the receiver structure designed based on a factor graph can be depicted using a receiver structure similar to Fig. 2(a), it is more convenient to represent the whole receiver using a factor graph as done, for example, in [111], [112], [116], and [117]. The sum-product algorithm is an iterative approach based on message-passing, and it can be applied to derive a wide variety of algorithms including the forward/backward algorithm, turbo decoding, and the Kalman filter [22].

Using the OFDM signal in (8) as an example, the optimum detector is

$$\hat{X}_{n,k} = \arg\max_{\{X_{n,k}\}} p(X_{n,k}|\mathbf{Y}) = \arg\max_{\{X_{n,k}\}} p(X_{n,k}, \mathbf{Y}) \qquad (32)$$

where $\mathbf{Y}$ is the received signal inside the observation window (e.g., one *transmit time interval* (TTI)), and the right-side equality follows from Bayes' theorem. For simplicity, consider only uncoded transmission here. If we take the channel response and other data symbols except the (n,k)th (denoted as $\mathbf{X}_{[n,k]}$) as intermediate variables, (32) can be rewritten as

$$\hat{X}_{n,k} = \arg\max_{\{X_{n,k}\}} \sum_{\mathbf{X}_{[n,k]}} \int_{\mathbf{H}} p(\mathbf{Y}, \mathbf{H}, \mathbf{X})\, d\mathbf{H}. \qquad (33)$$

The joint probability $p(\mathbf{Y}, \mathbf{H}, \mathbf{X})$ is exactly the global function in the framework of the sum-product algorithm. For simplicity, we consider here three subcarriers in a frequency-selective channel, denoted as H_1, H_2, H_3, respectively. Then, the factor graph is shown in Fig. 3(a).

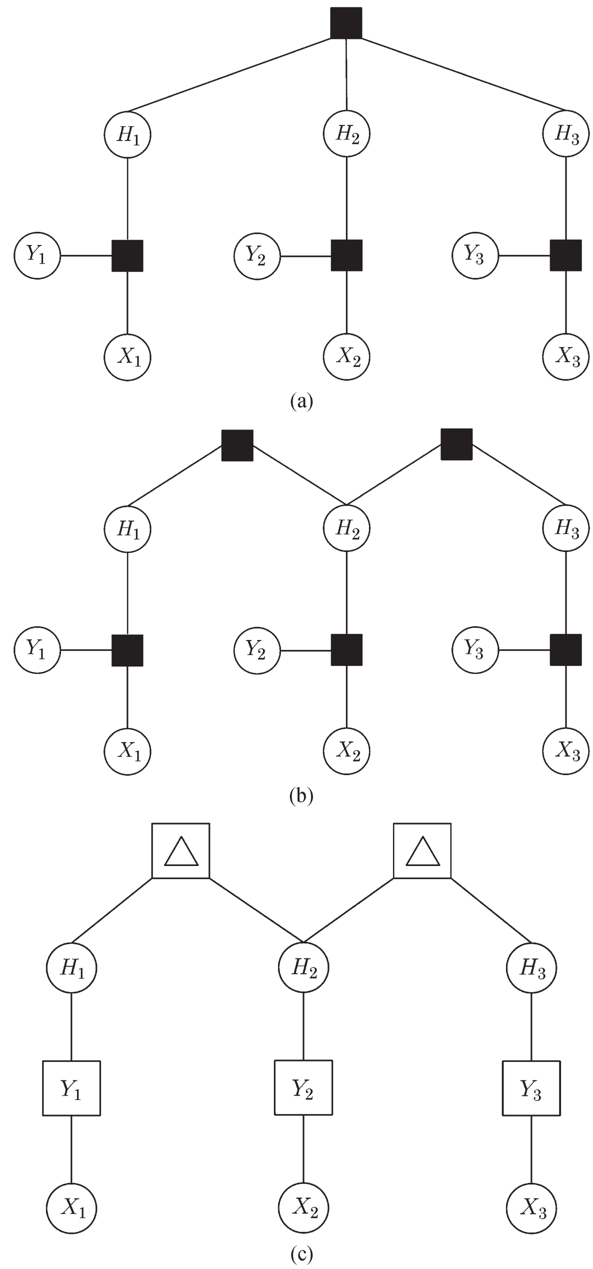


Fig. 3. (a) Factor graph for uncoded transmission. A black box indicates the check node and the circle indicates the variable node. (b) Factor graph for uncoded transmission in Markov channel. (c) Simplified factor graph for uncoded transmission in a Markov channel. A boxed Δ denotes a transfer node, a circled H_k denotes a channel response node, a boxed Y_k denotes an observation node, and a circled X_k denotes a data symbol node.

With message-passing, (33) can be solved in an iterative manner. When the factor graph is cycle-free, a node cannot receive the outgoing message sent by itself, and thus (33) can be solved exactly using the message update rule. Even though the cycles due to coded transmission cannot be avoided in the factor graph, short cycles that significantly degrade the estimation accuracy can be eliminated effectively given a sufficiently large code length and a good interleaver [22]. In the presence of cycles, the sum-product algorithm cannot lead to a unique schedule nor to a unique stopping criterion for the message passing. Various strategies for scheduling message passing are described in [118].

Note that the factor graph can also be represented in a Forney style [119]. Although the Forney style has several advantages [115] compared to the original ones in [22], we will use the original factor graph since it has been widely used in receiver design.

2) Channel Estimation: In a factor graph receiver, the purpose of channel estimation is to estimate the CFR, $H_{n,k}$. However, the CFR is invisible in (32) as it is only an intermediate variable. It becomes visible when the intermediate variable is taken into account as in (33). Due to the iterative property of message passing, the CFR can be estimated along with the message passing procedure among nodes in the graph. Since there is no *explicit* channel estimation stage in a factor graph receiver, the channel estimation can be considered as working *implicitly*. The iterative property of the sum-product algorithm implies that this implicit channel estimation also works in an iterative manner.

In nature, channel responses are correlated according to [120]. This increases the complexity in calculating the passed messages. To simplify the channel model, many works in factor graph receiver design, e.g., [111], [112], [116], [121], and [122], adopt a Markov assumption for the channel response. Taking the time direction for example, the channel response, H_n, is modeled by an auto-regressive model of first order [31]

$$H_{n+1} = \alpha H_n + v_n \tag{34}$$

where α is a coefficient that determines the correlation between successive fading coefficients, and v_k is a Gaussian distributed model input. In this case, we have

$$p(H_n|H_{n-1}, H_{n-2}, \cdots) = p(H_n|H_{n-1}) \tag{35}$$

and the channel now becomes a hidden Markov process since the channel response cannot be observed directly [22]. Due to this assumption, the global joint probability function can be further factored, and thus the factor graph in Fig. 3(a) can be further transformed to Fig. 3(b). In Fig. 3(b), node H_n is connected only to adjacent channel responses through check nodes; this greatly simplifies the calculation of the passed message.

3) Message Passing: Message passing is critical in a factor graph and the corresponding sum-product algorithm. For discrete variables, e.g., data symbol $X_{n,k}$, message is in the form of a list of probabilities in which $X_{n,k}$ takes points from a constellation, that is

$$\mu_{x\to f}(x) = \sum_m P_m \delta(x - x_m) \tag{36}$$

where $\mu_{x\to f}(x)$ denotes the message from variable node x to some check node f, and P_m is the probability of symbol x_m. When all the random variables under consideration are discrete variables, the message passing algorithm can work well. However, if there are continuous variables, e.g., channel response $H_{n,k}$, the messages are required to be in the form of probability density functions (pdf). Although the starting pdfs may be easily described, they generally become hard to manipulate after several iterations [123]. Hence, iterative channel estimation in this field is focused on how to pass a message that is in the form of a continuous pdf. This problem results from attempting to provide reliability information for channel estimates. Obviously, the continuous pdf can provide all information for the estimation results, including the reliability. Even if there is an approximation in practical message passing algorithms, the reliability information is always taken into account. In view of this, we can say that factor the graph receiver is more advanced than the turbo receiver, which can provide only hard estimates of the channel response.

Because it is difficult to pass messages in the form of continuous pdfs, many works employ canonical distributions to approximate the outgoing message of continuous variables

$$\mu_{x\to f}(x) = \sum_m \lambda_m c_m(\theta_m) \tag{37}$$

where $c_m(\theta_m)$ is the mth predetermined canonical distribution function with parameter θ_m, and λ_m is the corresponding coefficient. Canonical distributions provide a general framework for handling continuous variables. The problem of passing messages in the form of continuous pdfs is now reduced to passing only the parameters (i.e., λ_m's and θ_m's) so that the message receive node can recover the continuous pdfs approximately using the received parameters [23]. Canonical distributions can be considered as a general model for most practical message passing algorithms used in existing literature, e.g., discrete sampling in [124], and simpler heuristic techniques in [121].

Canonical distributions are first suggested in [23] as a framework for handling the outgoing messages for continuous variables. Several possible canonical distributions are discussed there as examples, but none are discussed in detail. In [122], [125], it is observed that the continuous pdf can be represented in an iterative form using the forward-backward algorithm [126]. Assuming that the forward-backward algorithm starts with a Gaussian pdf, the passed message becomes a mixture of more and more Gaussian-like pdfs as the message is being passed from node to node. Thus, it is natural to choose Gaussian pdfs as the canonical distributions. Then, the messages can be delivered by passing those parameters rather than the pdfs. For MPSK transmission, it is sufficient to only recover the channel phase. In [124] and [127], the channel phase is quantized into a finite number of values that correspond to letting the canonical distributions be a weighted sum of impulses. Then, the discrete forward-backward algorithm can be used to calculate the passed message. Other canonical distribution based approaches are also proposed in [123], [128], and [129] in the presence of unknown channel phase. In [128], the passed message is expanded using the Fourier series since the message pdf is a periodic function of the channel phase. When the channel is slowly time-varying, successive channel phases can be assumed to be the same; in this case, the solution is reduced to a sequence of Tikhonov pdfs [123].

Although the canonical distribution can approximate the continuous pdf well, it generally requires a significant computation load because there are many parameters to pass. A low complexity approach based on a transfer node is adopted in [111], [112], [116], and [117]. The transfer node is a new

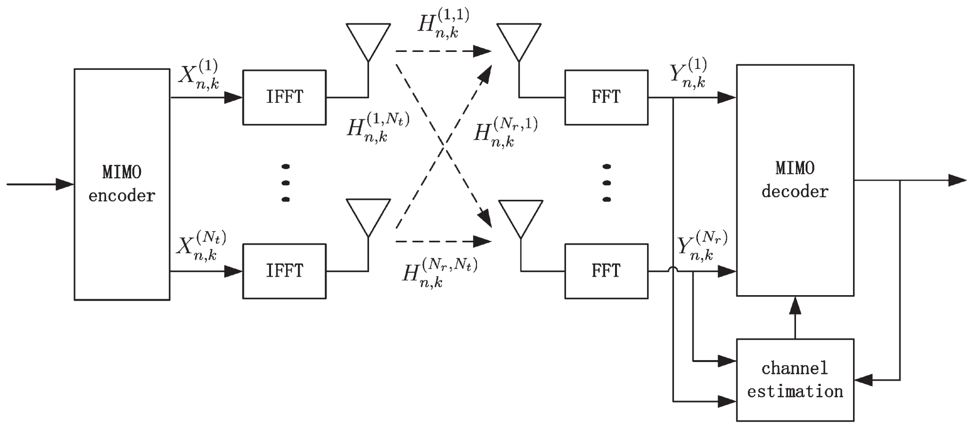


Fig. 4. MIMO-OFDM structure.

check node used in a factor graph to represent the variation between adjacent channel coefficients. Using a transfer node, the factor graph in Fig. 3(b) can be redrawn as in Fig. 3(c). Note that the factor graph in Fig. 3(c) is slightly different from the classical factor graph in [22] due to some simplification [111]. In Fig. 3(c), Δ is a random variable associated with each transfer node, and it is defined as the difference between adjacent channel responses, that is

$$\Delta = H_{n+1} - H_n. \tag{38}$$

In [111], Δ is treated as a zero-mean Gaussian random variable with variance σ_Δ^2 determined by the maximum Doppler shift in a flat fading channel. In this case, we have

$$\mathrm{E}(H_{n+1}) = \mathrm{E}(H_n), \tag{39}$$

$$\mathrm{Var}(H_{n+1}) = \mathrm{Var}(H_n) + \sigma_\Delta^2. \tag{40}$$

Since H_{n+1} is a complex Gaussian random variable, it is sufficient to pass only the mean and variance of H_n. The mean indicates the hard estimate for the current channel estimate, and the variance indicates the reliability. The transfer node based approach omits the structure of the message by treating it as a Gaussian random variable. Thus, it is sufficient to pass only two parameters to characterize the message pdf. Consequently, the complexity is reduced compared to canonical distribution based approaches. For OFDM transmission over successive symbols, the CFR is represented in a 2-D manner [8]. 2-D channel estimation using factor graphs are presented in [113] and [130], where Markov assumption for the channel responses is adopted in both the time and the frequency directions. The transfer node approach is also used in [113] and [130] to depict the correlation between adjacent channel coefficients in the same domain.

To avoid passing continuous pdfs as messages, a heuristic simplification is used in [121] for the outgoing message of continuous variables, where only the hard estimates of the channel responses are passed without reliability information. There are also works [131]–[133] resorting to variational message passing [134], [135] to approximate the continuous pdfs.

Different from the CFR-based and PM-based approaches, the ICE-based approach is not a stand-alone block in the receiver, but is a part of the whole iterative receiver. By taking into account the data symbols, the ICE-based approach has been proposed to improve the channel estimation performance of the CFR-based approach. In this sense, the ICE-based approach can be considered as an extension of traditional CFR-based channel estimation. Compared to the PM-based approach, it has no requirement on sparse channel parameters, and thus can be used in an urban environment. Besides, it has no need of long symbol sequence for averaging operation. Hence, it can easily be used in burst transmission systems. Despite these advantages, the major problem of the ICE-based approach is its complexity. In a turbo receiver, the demodulation-remodulation chain requires a significant amount of computation due to the iterative processing. Although there is no explicit iterative procedure in a factor-graph receiver, the calculation of the transmitted message between neighboring nodes can also result in a large computation load.

VI. CHANNEL ESTIMATION IN MIMO-OFDM

MIMO transmission greatly improves the capacity of wireless communications. Since OFDM can convert a frequency selective channel into parallel flat fading channels, it is natural to combine MIMO with OFDM to provide high rate data transmission over frequency selective channels. However, channel estimation in MIMO-OFDM systems is a challenging task due to the presence of multiple transmit antennas. Considering a MIMO-OFDM system with N_t transmit antennas and N_r receive antennas, as shown in Fig. 4, the signal at the jth receive antenna becomes

$$Y_{n,k}^{(j)} = \sum_{i=1}^{N_t} H_{n,k}^{(j,i)} X_{n,k}^{(i)} + W_{n,k}^{(j)} \tag{41}$$

where $H_{n,k}^{(j,i)}$ denotes the CFR from the ith transmit antenna to the jth receive antenna and $X_{n,k}^{(i)}$ denotes the signal transmitted from the ith transmit antenna. Because the received signal is now a superposition of signals from multiple transmit antennas, channel estimation approaches for SISO-OFDM system cannot be directly applied here. Hence, channel estimation in MIMO-OFDM systems needs to be reconsidered. We will discuss channel estimation approaches for MIMO-OFDM systems in this section.

A. CFR Based Channel Estimation

Strategies for channel estimation in MIMO-OFDM systems are different, depending on whether DD-CE or PA-CE is used. In DD-CE, initial channel estimation is accomplished using training symbols. One straightforward apporach is to send training symbols from only one transmit antenna during a given time interval while keeping the other transmit antennas silent [136]–[138]. As a result, the receive signal in (41) is reduced to the single antenna case in (8), and the channel response corresponding to each transmit antenna can be estimated using the approaches for SISO-OFDM systems. Assuming that the channels are quasi-static during the transmission for each antenna, the channel estimation performance can be improved if transmitting several training symbols during the training stage. In this case, the pilot subcarriers in the first training symbol are cyclically shifted to yield the second training symbol. This is equivalent to sampling the CFR at several different points [136], [139], [140]. Such a pilot arrangement can also improve the performance of channel estimation for the edge subcarriers since at least one subcarrier is at the edge of the available band.

Although the approaches above are straightforward, they have several drawbacks. First, the estimates of the CFRs on different antennas are obtained one by one. This increases the training stage, and thus reduces the spectral efficiency. Second, keeping silent increases the peak-average power ratio (PAPR), which is a key parameter for OFDM transmitter design [9]. Hence, transmitting the training symbols simultaneously is more preferred.

When transmitting the training symbols simultaneously, there are in total N_tK unknowns but only K equations are available for each receive antenna. Hence, rather than estimating the CFR directly, the receiver estimates the CIRs instead since the number of multipaths, L, is usually far less than the subcarrier number, K. Assuming that there are L paths with CIR, the total number of unknowns is reduced to N_tL. Given the received signal in (41), the estimates of the CIR from the ith transmit antenna to the jth receive antenna can be obtained by minimizing the following cost function:

$$\sum_{k=0}^{K-1}\left|Y_{n,k}^{(j)}-\sum_{i=0}^{N_t}\sum_{l=0}^{L-1}X_{n,k}^{(i)}h_{n,l}^{((j,i)}e^{-j\frac{2\pi kl}{K}}\right|^2. \tag{42}$$

Since the processing for each receive antenna is the same, we can drop the index of the receive antenna in what follows. Direct calculation of (42) yields:

$$\hat{\mathbf{h}}_n=\mathbf{Q}_n^{-1}\mathbf{p}_n \tag{43}$$

where

$$\hat{\mathbf{h}}_n=\left[h_{n,0}^{(i)},h_{n,1}^{(i)},\ldots,h_{n,L-1}^{(i)}\right]^{\mathrm{T}}$$

$$\mathbf{p}_n=\left[p_{n,0}^{(i)},p_{n,1}^{(i)},\ldots,p_{n,L-1}^{(i)}\right]^{\mathrm{T}}$$

$$\mathbf{Q}_n=\begin{bmatrix}\mathbf{Q}_{11} & \cdots & \mathbf{Q}_{1N_t}\\ \vdots & \ddots & \vdots\\ \mathbf{Q}_{N_t1} & \cdots & \mathbf{Q}_{N_tN_t}\end{bmatrix}$$

with

$$\mathbf{Q}_{(ij)}=\begin{bmatrix}q_{n,0}^{(ij)} & \cdots & q_{n,-L+1}^{(ij)}\\ \vdots & \ddots & \vdots\\ q_{n,L-1}^{(ij)} & \cdots & q_{n,0}^{(ij)}\end{bmatrix}$$

where

$$q_{n,l}^{(ij)}=\sum_{k=0}^{K-1}X_{n,k}^{(j)}X_{n,k}^{(i)*}e^{j\frac{2\pi kl}{K}}$$

and

$$p_{n,l}^{(i)}=\sum_{k=0}^{K-1}Y_{n,k}X_{n,k}^{(i)*}e^{j\frac{2\pi kl}{K}}.$$

The approach above is first introduced in [25]. Note that, there is no particular design for the training symbols; so, this approach is also applicable when the training symbols are generated using the decoded data symbols. Matrix inversion in (43) requires a lot of computational complexity. Although significant tap catch is adopted in [25] to reduce the complexity required by matrix inversion, it is still very complicated. In [26], optimum training sequences are proposed such that $q_{n,l}^{(ii)}=K\delta[l]$ where $\delta[\cdot]$ is the unit impulse function. In this case, $\mathbf{Q}_n$ in (43) becomes a diagonal matrix, and thus the complexity is greatly reduced. In [26], a simplification, based on recursive strategies, is used for channel estimation when the system is in data mode. Note that the training symbol design in [26] can only be used in situations where the number of transmit antennas is less than K/L.

Assuming that channel is quasi-static over two sequential OFDM symbols, it is also possible to reduce the complexity by using space-time block coded (STBC) training symbols. In [141] and [142], training symbols from multiple transmit antennas are designed based on Alamouti's scheme [143]. As a result, CFR estimation in the receiver can be carried out without matrix inversion; however, when the channel is not constant over a number of sequential OFDM symbols equal to the number of transmit antennas, the performance will be significantly degraded.

Besides reducing the complexity, the orthogonal design of training symbols can also achieve the optimum MSE performance [26], [56]. Other sequences can also be used for the same purpose [4], [144], [145]. In [144], the Golay code is found to have lower PAPR, and the Zadoff-Chu sequence [145] can provide constant amplitude. A general framework for optimum training signal design is proposed in [146], which includes all existing training signals as special cases.

For PA-CE, the silent pilot approach is widely used in MIMO-OFDM based wireless systems [4]. Hence, PA-CE in MIMO-OFDM systems can directly use the approaches developed for SISO-OFDM at the expense of reduced channel sampling density. Optimum pilot design for PA-CE in MIMO-OFDM has been investigated in [147]. The technique used for achieving optimum MSE performance in [147] is also adopted in [56] and [148] for optimum pilot design in SISO-OFDM. As shown in [147], for a single OFDM symbol, the pilots should be allocated with equal power and equally spaced in the frequency domain; and pilot sequences on different transmit antennas should be orthogonal, which can be achieved by shifting the phases of those sequences. For multiple OFDM symbols, there should be constant shifts between the indices of the frequency domain pilots for the current OFDM symbol and those for the next OFDM symbol.

B. PM-Based Channel Estimation

PM-based channel estimation can also be used in MIMO-OFDM systems. By allowing silent pilots for multiple transmit antennas, the approach proposed in [13] can be directly used in MIMO-OFDM systems for estimating path delays, at the cost of reduced sampling density in the frequency domain. In [149], it is found that phase orthogonal sequences can convert a multiple-antenna case into a single-antenna one, and, thus, the ESPRIT algorithm can be used to estimate the path delays [13]. Another near-ML path delay estimation approach, also presented in [149], is similar to the alternating maximization proposed in [150]. As shown in [150], alternating maximization works in an iterative manner: in each iteration, a maximization is performed with respect to a single parameter while the other parameters remain unchanged. In [151], by assuming that the actual path delays take values from a given set, the initial estimation of the path delays is converted into detection of the path locations. *Probabilistic data association* (PDA), which is first proposed in [152] for multiuser detection in *code-division multiple access* (CDMA) systems, is adopted in [151] to detect the path delays. Since the predetermined delays are different from the actual path delays, decision feedback is used in [151] to refine the path delay estimation in an iterative manner.

It is also possible to exploit the sparsity of the channel in sample-spaced CIR. In [78], a *matching pursuit* (MP) algorithm is used to determine the positions of the *most significant taps* (MST), as we have introduced in Section V. The MP algorithm is also adopted in [153] for sparse channel recovery. In [154], the positions of the MST are determined by exploiting the second-order statistics of the receive signal. Note that the second-order statistics can also be utilized for blind channel estimation in MIMO-OFDM systems, e.g., [155] and [156].

C. Iterative Channel Estimation

Iterative channel estimation is also used in MIMO-OFDM. Depending on the framework of receiver structure, iterative channel estimation is modified to adapt MIMO-OFDM transmission.

For a turbo receiver, iterative channel estimation in MIMO-OFDM is different from that in SISO-OFDM in two ways. First, a soft MIMO detector is required in the iterative loop. One intuitive way to construct a soft MIMO detector is to use a traditional zero-forcing or LMMSE equalizer [28] followed by a group of soft demappers [41] for each equalized stream. Although this approach is straightforward, the linear equalizers suffer significant performance degradation when correlation exists between antennas [157]. Instead, one can use a *list sphere decoder* (LSD) [158] to obtain a low complexity soft MIMO detector. The LSD can be considered as the soft output version of the sphere detector introduced in [159] and [160]. It also has several variants (see [161] for more detailed information). Second, unlike pilot symbols that can be silent when another transmit antenna is emitting, data symbols are always transmitted simultaneously. In this case, the iterative CFR estimation used in SISO-OFDM is not applicable even through these data symbols are known. One way to overcome this problem is to estimate the CIR instead of the CFR. By treating the data symbols as known training symbols, the approach proposed in [26] can be used to estimate the CIR for each receive antenna. It is also possible to estimate the CIR with the *space-alternating expectation-maximization* (SAGE) algorithm [84]. SAGE is in essence an EM-based algorithm but can converge faster than the traditional EM algorithm.

For factor graph receivers, the approach for channel estimation in MIMO-OFDM is basically the same as for SISO-OFDM, that is, passing messages among nodes in the factor graph. However, the factor graph representing a MIMO-OFDM receiver is different from that of a SISO-OFDM receiver because multiple receive antennas must be taken into account. A MIMO system has been considered in [116] and [117] for a factor graph receiver. In a MIMO-OFDM system, since there are more variables to represent in the factor graph, a *three-dimensional* (3-D) factor graph may be required to represent the overall receiver [113], [130], [162]. In this case, it is difficult to use complex message passing algorithms, e.g., canonical distribution. Hence, many works on factor graph receivers for MIMO-OFDM reception resort to the transfer node approach to reduce complexity. Using channel correlation, the CFR is transformed into the CIR in a turbo receiver in order to reduce the number of unknowns. In a factor graph receiver, however, there is no need to make this transformation. CFR estimates can be directly obtained along with message passing. At a first glance, it seems surprising because there are more unknowns than equations in the receiver. However, if we observe the factor graph, e.g., Fig. 3(b), we find that a particular channel response node gathers messages from the connected receive signal node and other neighboring channel response nodes to yield its own channel estimate. This means that the channel correlation has been implicitly utilized.

VII. Channel Estimation in Practical Systems

Along with the wide adoption of OFDM transmission in current commercial systems, channel estimation techniques have been practically applied in existing systems [163]. In this section, we will make a brief introduction about the channel estimation of OFDM transmission in practical systems. The application in *Long-Term Evolution* (LTE) [4] will be emphasized since it is the most popular cellular system today.

OFDM is adopted for downlink transmission in LTE. The resource allocation in downlink transmission is carried out in units of *resource block*s (RB). For a normal length CP, each RB consists of twelve subcarriers and seven OFDM symbols. Four pilot symbols are inserted into each RB for channel estimation, as shown in Fig. 5(a). Obviously, a typical way for channel estimation in LTE downlink is to use the PA-CE in Section III. The LS estimates corresponding to the pilot symbols are obtained

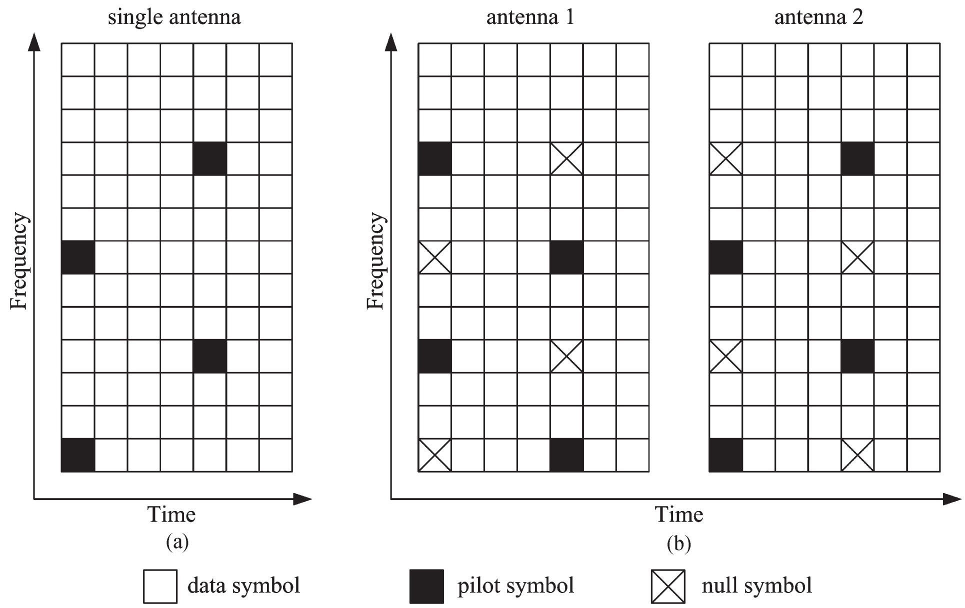


Fig. 5. Pilot pattern in LTE donwlink transmission (a) for single transmit antenna (b) for two transmit antennas.

first, and then different interpolation approaches can be used to obtain the channel estimates corresponding to the data symbols. Channel estimation performance can be further improved if we adopt the iterative channel estimation in Section V. LTE also permits the use of use multiple transmit antennas to improve the system performance. To estimate the channel from multiple transmit antennas, LTE requires that only one antenna be allowed to transmit at a time while the other transmit antennas remain silent. Taking two transmit antennas, for example, the pilot pattern in this case is shown in Fig. 5(b). As we can see, a null symbol is inserted for one antenna when the other one is transmitting.

For uplink transmission in LTE, SC-FDMA is adopted to reduce the PAPR of the transmit signals. SC-FDMA in LTE can be implemented using DFT-spread OFDM, and, thus, the structure of the transceiver for SC-FDMA is similar to that for OFDM except for DFT spreading. Although it is possible to multiplex data symbols and pilots in one symbol duration [93], LTE adopts a simple block-type pilot pattern [164] for uplink transmission, that is, only the pilot is transmitted in one symbol duration. In the LTE standard, the pilot is placed at the fourth symbol in one slot for a normal CP length.

Channel estimation approaches are also applied in other systems employing OFDM transmission. In IEEE 802.16m-based *Worldwide Interoperability for Microwave Access* (WiMAX) networks [5], pilot symbols are multiplexed with data symbols. This is similar to the pilot pattern in LTE, and thus PA-CE based channel estimation approaches can be used. In IEEE 802.11a based-*Wireless Local Area Network* (WLAN) [6], training symbols are first sent at the beginning of each frame, and then comb-type pilots are multiplexed with data symbols to track the time variation of channel.

VIII. Open Issues and Future Work

In the above, we have discussed channel estimation for OFDM systems. Iterative channel estimation is emphasized since it can greatly improve the system performance. Three open issues can be found:

1) The MSE performance achieved using the LMMSE estimator has been regarded as the MSE bound for conventional channel estimation. Since the receiver now works in an iterative manner, the performance of channel estimation should be improved as the iteration proceeds. It is therefore useful to determine a performance bound for iterative channel estimation.
2) Another question is how to exploit soft channel estimation in a turbo receiver to provide reliability information about the channel estimate rather than only the hard estimate. In this case, the performance of channel estimation in a turbo receiver should be improved.
3) Factor graph, along with the Bayesian network, provides a graph model based receiver design. It is interesting to compare the factor graph receiver and the turbo receiver (with and without reliability information of channel estimation). Moreover, traditional analysis only provides a satisfactory explanation when the graph is cycle-free. Although EXIT chart or density evolution can be used to explain the performance of the receiver in [103], [105], [165] and [166], there is still no analytical explanation on the behavior of the factor graph receiver when there are loops in the graph [115], [126].

IX. Summary

In this survey, we have reviewed channel estimation for OFDM systems. Four topics have been covered: traditional channel estimation based on the CFR, PM based channel estimation, iterative channel estimation, and channel estimation in MIMO-OFDM systems. As an emerging receiver structure, the iterative receiver structure can greatly improve the system performance. Hence, in this survey, we emphasized iterative channel estimation for a turbo receiver and for a factor graph receiver. Other types of channel estimation approaches have also been reviewed to provide a complete picture on this topic.

References

[1] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency division multiplexing using the discrete Fourier transform," *IEEE Trans. Commun. Technol.*, vol. COM-19, no. 5, pp. 628–634, Oct. 1971.

[2] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, no. 5, pp. 5–14, May 1990.

[3] L. J. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.*, vol. COM-33, no. 4, pp. 666–675, Jul. 1985.

[4] *Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation*, 3GPP Std. Rev. 36.211 (V9), Sep. 2008.

[5] *IEEE Standard for Local and Metropolitan Area Neworks*, IEEE Std. Rev. 802.16m, 2011.

[6] *IEEE 802.11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Std. Rev. 2012, 2012.

[7] Y. G. Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 902–915, Jul. 1998.

[8] Y. G. Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 49, no. 4, pp. 1207–1215, Jul. 2000.

[9] T. Hwang, C. Yang, G. Wu, S. Li, and G. Y. Li, "OFDM and its wireless applications: A survey," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1673–1694, May 2009.

[10] M. K. Ozdemir and H. Arslan, "Channel estimation for wireless OFDM systems," *IEEE Commun. Surveys Tuts.*, vol. 9, no. 2, pp. 18–49, 2007.

[11] J. J. Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Borjesson, "On channel estimation in OFDM systems," in *Proc. IEEE VTC*, Jul. 1995, vol. 2, pp. 815–819.

[12] O. Edfors, M. Sandell, J. J. Beek, S. K. Wilson, and P. O. Borjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 931–939, Jul. 1998.

[13] B. Yang, K. B. Letaief, R. S. Cheng, and Z. Cao, "Channel estimation for OFDM transmission in multipath fading channels based on parametric channel modeling," *IEEE Trans. Commun.*, vol. 49, no. 3, pp. 467–479, Mar. 2001.

[14] J. T. Chen, A. Paulraj, and U. Reddy, "Multicahnnel maximum-likehood sequence estimation (MLSE) equalizer for GSM using a parametric channel model," *IEEE Trans. Commun.*, vol. 47, no. 1, pp. 53–63, Jan. 1999.

[15] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *Conf. Rec. IEEE ICC*, May 1993, pp. 1064–1070.

[16] F. Sanzi, S. Jelting, and J. Speidel, "A comparative study of iterative channel estimators for mobile OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 849–859, Sep. 2003.

[17] R. Gallager, "Low-density parity-check codes," *IRE Trans. Inf. Theory*, vol. 8, no. 1, pp. 21–28, Jan. 1962.

[18] D. Mackay, "Good error correcting codes based on very sparse matrices," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 399–431, Mar. 1999.

[19] N. Alon and M. Luby, "A linear time erasure-resilient code with nearly optimal recovery," *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 1732–1736, Nov. 1996.

[20] J. R. Barry, Low-Desntiy Parity-Check Codes. [Online]. Available: http://www.sce.carleton.ca/hsaeedi/ldpc.pdf

[21] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 5, pp. 533–547, Sep. 1981.

[22] F. R. Kschischange, B. J. Frey, and H. A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 498–519, Feb. 2001.

[23] A. P. Worthen and W. E. Stark, "Unified design of iterative receivers using factor graph," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 843–849, Feb. 2001.

[24] G. J. Foschini, "Layered space-time architecture for wireless communicatin in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[25] G. Y. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 461–471, Mar. 1999.

[26] G. Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 67–75, Jan. 2002.

[27] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Pilot-based channel estimation for OFDM systems by tracking the delay-subspace," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 315–325, Jan. 2004.

[28] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[29] Y. G. Li and L. J. Cimini, "Bounds on the interchannel interference of OFDM in time-varying impairments," *IEEE Trans. Commun.*, vol. 49, no. 3, pp. 401–404, Mar. 2001.

[30] N. Benvenuto, R. Dinis, D. Falconer, and S. Tomasin, "Single carrier modulation with nonlinear frequency domain equalization: An idea whose time has come-again," *Proc. IEEE*, vol. 98, no. 1, pp. 69–96, Jan. 2010.

[31] S. Haykin, *Adaptive Filter Theory (Fourth Edition)*. Beijing, China: Publishing House of Electronics Industry, 2010.

[32] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

[33] N. Chen and G. T. Zhou, "What is the price paid for superimposed training in OFDM?" in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2004, vol. 4, pp. 421–424.

[34] G. T. Zhou and N. Chen, "Superimposed training for doubly seletive channels," in *Proc. IEEE Workshop Stat. Signal Process.*, Sep. 2003, pp. 82–85.

[35] P. Hoeher and F. Tufvesson, "Channel estimation with superimposed pilot sequence," in *Proc. IEEE GLOBECOM*, Dec. 1999, vol. 4, pp. 2162–2166.

[36] J. K. Tugnait and W. Luo, "On channel estimation using superimposed training and first-order statistics," *IEEE Commun. Lett.*, vol. 7, no. 9, pp. 413–415, Sep. 2003.

[37] A. G. Orozco-Lugo, M. M. Lara, and D. C. McLernon, "Channel estimation using implicit training," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 240–254, Jan. 2004.

[38] N. Chen and G. T. Zhou, "Superimposed training for OFDM: A peak-to-average power ratio analysis," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2277–2287, Jun. 2006.

[39] V. Mignone, A. Morello, and M. Visintin, "CD3-OFDM: A new channel estimation method to improve the spectrum efficiency in digital terrestrial television systems," in *Proc. IBC*, Sep. 1995, pp. 122–128.

[40] S. Y. Park, Y. G. Kim, and C. G. Kang, "Iterative receiver for joint detection and channel estimation in OFDM systems under mobile radio channels," *IEEE Trans. Veh. Technol.*, vol. 53, no. 2, pp. 450–460, Mar. 2004.

[41] S. T. Brink, J. Speidel, and R. H. Yan, "Iterative demapping for QPSK modulation," *Electron. Lett.*, vol. 34, no. 15, pp. 1459–1460, Jul. 1998.

[42] J. Zhang, X. Mu, E. Chen, and S. Yang, "Decision-directed channel estimation based on iterative linear minimum mean square error for orthogonal frequency division multiplexing systems," *IET Commun.*, vol. 3, no. 7, pp. 1136–1143, Jul. 2009.

[43] Q. Sun, D. C. Cox, H. C. Huang, and A. Lozano, "Estimation of continuous flat fading MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 549–553, Oct. 2002.

[44] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Trans. Broadcast.*, vol. 48, no. 3, pp. 223–229, Sep. 2002.

[45] S. Kalyani and K. Giridhar, "Mitigation of error propagation in decision directed OFDM channel tracking using generalized M estimators," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1659–1672, May 2007.

[46] M. Sandell and O. Edfors, "A comparative study of pilot-based channel estimators for wireless OFDM." [Online]. Available: http://pure.ltu.se/portal/files/1614570/Report.pdf

[47] P. Hoeher, S. Kaiser, and P. Robertson, "Two-dimensional pilot-symbol-aided channel estimation by Wiener filtering," in *Proc. IEEE ICASSP*, Apr. 1997, vol. 3, pp. 1845–1848.

[48] S. G. Kang, Y. M. Ha, and E. K. Joo, "A comparative investigation on channel estimatin algorithms for OFDM in mobile communications," *IEEE Trans. Broadcast.*, vol. 49, no. 2, pp. 142–149, Jun. 2003.

[49] A. Dowler and A. Nix, "Performance evaluation of channel estimation techniques in a multiple antenna OFDM systems," in *Proc. IEEE VTC*, Orlando, FL, USA, Oct. 2003, vol. 2, pp. 1214–1218.

[50] S. Coleri, M. Ergen, A. Pri, and A. Bahai, "A study of channel estimation in OFDM systems," in *Proc. IEEE VTC*, Vancouver, BC, Canada, Sep. 2002, pp. 894–898.

[51] X. Wang and K. J. R. Liu, "OFDM channel estimation based on time-frequency polynomial model of fading multipath channel," in *Proc. IEEE VTC*, Atlantic City, NJ, USA, Oct. 2001, vol. 1, pp. 460–464.

[52] M. X. Chang and Y. T. Su, "Model-based channel estimation for OFDM signals in Rayleigh fading," *IEEE Trans. Commun.*, vol. 50, no. 4, pp. 540–544, Apr. 2002.

[53] P. Hoeher, S. Kaiser, and P. Robertson, "Pilot-symbol-aided channel estimation in time and frequency," in *Proc. Int. Workshop Multi-Carrier Spread-Spectrum*, Apr. 1997, pp. 169–178.

[54] M. H. Hsieh and C. H. Wei, "Channel estimation for OFDM systems based on comb-type pilot arrangement in frequency selective fading channels," *IEEE Trans. Consum. Electron.*, vol. 44, no. 1, pp. 217–225, Feb. 1998.

[55] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3065–3073, Dec. 2001.

[56] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Trans. Consum. Electron.*, vol. 44, no. 3, pp. 1122–1128, Aug. 1998.

[57] J. W. Choi and Y. H. Lee, "Optimum pilot pattern for channel estimation in OFDM systems," *IEEE Trans Wireless Commun.*, vol. 4, no. 5, pp. 2083–2088, Sep. 2005.

[58] X. Cai and G. B. Giannakis, "Error probability minimization pilots for OFDM with M-PSK modulaion over Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 146–155, Jan. 2004.

[59] S. Ohno and G. B. Giannakis, "Capacity maximizing MMSE-optimal pilots for wireless OFDM over frequency-selective block Rayleigh-fading channels," *IEEE Trans. Inf. Theoy*, vol. 50, no. 9, pp. 2138–2345, Sep. 2004.

[60] X. Meng and J. K. Tugnait, "Semi-blind channel estimation and detection using superimposed training," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Montreal, BC, Canada, May 2004, vol. 4, pp. 417–420.

[61] H. Z. Jafarian, M. J. Omidi, and S. Pasupathy, "Improved channel estimation using noise reduction for OFDM systems," in *Proc. IEEE VTC*, Jeju, Korea, Apr. 2003, vol. 2, pp. 1308–1312.

[62] Y. H. Yeh and S. G. Chen, "Efficient channel estimation based on discrete cosine transform," in *Proc. IEEE ICASSP*, Hong Kong, Apr. 2000, vol. 4, pp. 676–679.

[63] Y. H. Yeh and S. G. Chen, "DCT-based channel estimation for OFDM systems," in *Proc. IEEE ICC*, Paris, France, Jun. 2004, vol. 4, pp. 2442–2446.

[64] M. Diallo, R. Rabineau, and L. Cariou, "Robust DCT based channel estimation for MIMO-OFDM systems," in *Proc. IEEE WCNC*, Apr. 2009, pp. 1–5.

[65] Y. Zhao and A. Huang, "A novel channel estimation method for OFDM mobile communicatin systems based on pilot signals and transform domain processing," in *Proc. IEEE VTC*, May 1997, vol. 3, pp. 2089–2093.

[66] A. L. Swindlehurst, "Time delay and spatial signal estimation using known asynchronous signals," *IEEE Trans. Signal Process.*, vol. 46, no. 2, pp. 449–462, Feb. 1998.

[67] R. R. T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.

[68] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1996.

[69] P. Stoica and R. Moses, *Introduction to Spectral Analysis*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1997.

[70] M. R. Raghavendra, E. Lior, S. Bhashyam, and K. Giridhar, "Parametric channel estimation for pseudo-random tile-allocation in uplink OFDMA," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5370–5381, Nov. 2007.

[71] P. Strobach, "Low-rank adaptive filters," *IEEE Trans. Signal Process.*, vol. 44, no. 12, pp. 2932–2947, Dec. 1996.

[72] M. R. Raghavendra, S. Bhashyam, and K. Giridhar, "Exploiting hopping pilots for parametric channel estimation in OFDM systems," *IEEE Signal Process. Lett.*, vol. 12, no. 11, pp. 737–740, Nov. 2005.

[73] I. C. Wong and B. L. Evans, "Sinusoidal modeling and adaptive channel prediction in mobile OFDM systems," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1601–1615, Apr. 2008.

[74] G. Xu, I. R. H. Roy, and T. Kailath, "Detection of number of sources via exploitation of centro-symmetro property," *IEEE Trans. Signal Process.*, vol. 42, no. 1, pp. 102–112, Jan. 1994.

[75] H. Li, D. Liu, J. Li, and P. Stoica, "Channel order and RMS delay spread estimation with application to AC power line communications," *Digit. Signal Process., Rev. J.*, vol. 13, no. 2, pp. 284–399, Apr. 2003.

[76] M. R. Raghavendra and K. Giridhar, "Improving channel estimation in OFDM systems for sparse multipath channels," *IEEE Signal Process. Lett.*, vol. 12, no. 1, pp. 52–55, Jan. 2005.

[77] M. R. Raghavendra, S. Bhashyam, and K. Giridhar, "Improving channel estimation in OFDM systems for sparse multipath channnels," in *Proc. IEEE Signal Process. Adv. Wireless Commun.*, Jul. 2004, pp. 106–109.

[78] D. Wang, B. Han, J. Zhao, X. Gao, and X. You, "Channel estimation algorithms for broadband MIMO-OFDM sparse channel," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun.*, 2003, pp. 1929–1933.

[79] S. F. Cotter and B. D. Rao, "Sparse channel estimation via matching pursuit with application to equalization," *IEEE Trans. Commun.*, vol. 50, no. 3, pp. 374–377, Mar. 2002.

[80] J. Bonnet and G. Auer, "Optimized iterative channel estimation for OFDM," in *Proc. IEEE VTC-Fall*, 2006, pp. 1–5.

[81] Y. N. Lee, A. Ashikhmin, and J. T. Chen, "Impact of soft channel construction on iterative channel estimation and data decoding for multi-carrier systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2762–2770, Jul. 2008.

[82] T. Kang and R. A. Iltis, "Iterative carrier frequency offset and channel estimation for underwater acoustic OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 9, pp. 1650–1661, Dec. 2008.

[83] D. Yoon and J. Moon, "Low-complexity iterative channel estimation for turbo receivers," *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 1182–1187, May 2012.

[84] J. Ylioinas and M. Juntti, "Iterative joint detection, decoding, and channel estimation in turbo-coded MIMO-OFDM," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1784–1796, May 2009.

[85] J. Hagenauer, "The Turbo principle: Tutorial introduction and state of the art," in *Proc. Symp. Turbo Codes*, Brest, France, Sep. 1997, pp. 1–11.

[86] L. Hanzo, T. H. Liew, and B. L. Yeap, *Turbo Coding, Turbo Equalisation and Space-Time Coding*. Hoboken, NJ, USA: Wiley, 2002.

[87] M. Zhao, Z. Shi, and M. C. Reed, "Iterative turbo channel estimation for OFDM system over rapid dispersive fading channel," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3173–3184, Aug. 2008.

[88] M. C. Valenti and B. D. Woerner, "Refined channel estimation for coherent detection of turbo codes over flat-fading channels," *Electron. Lett.*, vol. 34, no. 17, pp. 1648–1649, Aug. 1998.

[89] M. C. Valenti and B. D. Woerner, "Iterative channel estimation and decoding of pilot symbol assisted turbo codes over flat-fading channels," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 9, pp. 1697–1705, Sep. 2001.

[90] H. J. Su and E. Geraniotis, "Low-complexity joint channel estimation and decoding for pilot symbol-assisted modulation and multiple differential detection systems with correlated Rayleigh fading," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 249–261, Feb. 2002.

[91] B. L. Yeap, C. H. Wong, and L. Hanzo, "Reduced complexity in-phase/quadrature-phase M-QAM turbo equalization using iterative channel estimation," *IEEE Trans. Wireless Commun.*, vol. 2, no. 1, pp. 2–10, Jan. 2003.

[92] Y. Huang and J. A. Ritcey, "Joint iterative channel estimation and decoding for bit-interleaved coded modulation over correlated fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2549–2558, Sep. 2005.

[93] C. T. Lam, D. D. Falconer, and F. D. Lemoine, "Iterative frequency domain channel estimation for DFT-precoded OFDM systems using in-band pilots," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 2, pp. 348–358, Feb. 2008.

[94] C. T. Lam, D. D. Falconer, and F. D. Lemoine, "A low complexity frequency domain iterative decision-directed channel estimation technique for single-carrier systems," in *Proc. IEEE VTC*, 2007, pp. 1966–1970.

[95] Y. Liu and S. Sezginer, "Two iterative channel estimation algorithms in single-input multiple-output (SIMO) LTE systems," *Trans. Emerging Telecommun. Technol.*, vol. 24, no. 1, pp. 59–68, Jan. 2013.

[96] H. Niu and J. A. Ritcey, "Joint iterative channel estimation and decoding of pilot symbol assisted LDPC coded QAM over flat fading channels," in *Proc. 36th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2003, vol. 2, pp. 2265–2269.

[97] G. Auer and J. Bonnet, "Threshold controlled iterative channel estimation for coded OFDM," in *Proc. IEEE VTC*, Apr. 2007, pp. 1737–1741.

[98] D. Kim, H. M. Kim, and G. H. Im, "Iterative channel estimation with frequency replacement for SC-FDMA systems," *IEEE Trans. Commun.*, vol. 60, no. 7, pp. 1877–1888, Jul. 2012.

[99] E. Jaffrot and M. Siala, "Turbo channel estimation for OFDM systems on highly time and frequency selective channels," in *Proc. IEEE ICASSP*, 2000, vol. 5, pp. 2977–2980.

[100] X. Zhuang and F. W. Vook, "Iterative channel estimation and decoding for a turbo-coded OFDM system via EM algorithm," in *Proc. IEEE ICASSP*, May 2002, vol. 3, pp. 2337–2340.

[101] B. Lu, X. Wang, and Y. G. Li, "Iterative receivers for space-time block coded OFDM systems in dispersive fading channels," *IEEE Trans. Wireless Commun.*, vol. 1, no. 2, pp. 213–225, Apr. 2002.

[102] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood estimation from incomplete data," *J. Roy. Stat. Soc. B, Methodol.*, vol. 39, no. 1, pp. 1–38, 1977.
[103] S. T. Brink, "Convergence of iterative decoding," *Electron. Lett.*, vol. 35, no. 10, pp. 806–808, May 1999.
[104] S. T. Brink, "Iterative decoding for multicode CDMA," in *Proc. IEEE VTC*, May 1999, pp. 1876–1880.
[105] S. T. Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.
[106] J. S. Liu, *Monte Carlo Strategies in Scientific Computing*. New York, NY, USA: Springer-Verlag, 2003.
[107] R. A. Levine and G. Casella, "Implementations of the Monte Carlo EM algorithm," *J. Comput. Graph. Stat.*, vol. 10, no. 3, pp. 422–439, Sep. 2001.
[108] S. Song and K. M. Sung, "Soft input channel estimation for turbo equalization," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2885–2894, Oct. 2004.
[109] X. Wang and R. Chen, "Blind turbo equalization in Gaussian and impulsive noise," *IEEE Trans. Veh. Technol.*, vol. 50, pp. 1092–1105, Jul. 2001.
[110] R. Chen, J. S. Liu, and X. Wang, "Convergence analysis and comparisons of Markov chain Monte Carlo algorithms in digital communications," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 255–270, Feb. 2002.
[111] T. Wo, P. A. Hoeher, and Z. Shi, "Graph-based soft channel estimation for fast fading channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4243–4251, Dec. 2012.
[112] T. Wo, J. C. Fricke, and P. A. Hoeher, "A graph-based iterative Gaussian detector for frequency-selective MIMO channels," in *Proc. IEEE ITW*, Chengdu, China, Oct. 2006, pp. 581–585.
[113] C. Knievel, Z. Shi, P. A. Hoeher, and G. Auer, "2D graph-based soft channel estimation for MIMO-OFDM," in *Proc. IEEE ICC*, May 2010, pp. 1–5.
[114] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, CA, USA: Morgan Kaufmann, 1988.
[115] H. A. Loeliger, "An introduction to factor graph," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 28–41, Jan. 2004.
[116] T. Wo, C. Liu, and P. A. Hoeher, "Graph-based iterative Gaussian detection with soft channel estimation for MIMO systems," in *Proc. 7th ITG Conf. SCC*, Jan. 2008, pp. 1–6.
[117] T. Wo, C. Liu, and P. A. Hoeher, "Graph-based soft channel and data estimation for MIMO systems with asymmetric LDPC codes," in *Proc. IEEE ICC*, May 2008, pp. 620–624.
[118] D. Fertonani, A. Barbieri, and G. Colavolpe, "Novel graph-based algorithms for soft-output detection over dispersive channels," in *Proc. IEEE GLOBECOM*, 2008, pp. 1127–1131.
[119] G. D. Forney, "Codes on graphs: Normal realizations," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 520–548, Feb. 2001.
[120] R. H. Clarke, "A statistical theory of mobile radio reception," *Bell Syst. Tech. J.*, vol. 47, no. 6, pp. 957–1000, Jul./Aug. 1968.
[121] H. Niu, M. Shen, J. A. Ritcey, and H. Liu, "A factor graph approach to iterative channel estimation and LDPC decoding over fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1345–1350, Jul. 2005.
[122] Y. Zhu, D. Guo, and M. L. Honig, "A message-passing approach for joint channel estimation, interference mitigation, decoding," *IEEE Trans. Wireless Commmun.*, vol. 8, no. 12, pp. 6008–6018, Dec. 2009.
[123] G. Colavolpe, A. Barbieri, and G. Caire, "Algorithms for iterative decoding in the presence of strong phase noise," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 9, pp. 1748–1757, Sep. 2005.
[124] C. Komninakis and R. D. Wesel, "Joint iterative channel estimation and decoding in flat correlated Rayleigh fading," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 9, pp. 1706–1716, Sep. 2001.
[125] G. Ferrari, G. Colavolpe, and R. Raheli, *Detection Algorithms for Wireless Communications*. New York, NY, USA: Wiley, 2004.
[126] R. J. McEliece, D. J. C. MacKay, and J. F. Cheng, "Turbo decoding as an instance of Pearl's 'belief propagation' algorithm," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 2, pp. 140–152, Feb. 1998.
[127] J. Dauwels and H. A. loeliger, "Joint decoding and phase estimation: An exercise in factor graphs," in *Proc. IEEE Symp. Inf. Theory*, Yokohama, Japan, Jul. 2003, p. 231.
[128] A. Barbieri, G. Colavolpe, and G. Caire, "Joint iterative detection and decoding in the presence of phase noise and frequency offset," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 171–179, Jan. 2007.
[129] J. Dauwels and H. A. Loeliger, "Phase estimation by message passing," in *Proc. IEEE ICC*, Paris, France, Jun. 2004, pp. 523–527.
[130] C. Knievel, P. A. Hoeher, A. Tyrrell, and G. Auer, "Multi-dimensional graph-based soft iterative receive for MIMO-OFDM," *IEEE Trans. Commun.*, vol. 60, no. 6, pp. 1599–1609, Jun. 2012.
[131] G. E. Kirkelund, C. N. Manchon, L. P. B. Christensen, E. Riegler, and B. H. Fleury, "Variational message-passing for joint channel estimation and decoding in MIMO-OFDM," in *Proc. IEEE GLOBECOM*, 2010, pp. 1–6.
[132] H. Attias, "A variational Bayesian framework for graphical models," in *Proc. Adv. Neural Inf. Process. Syst.*, 2000, pp. 209–215.
[133] M. Beal, "Variational algorithms for approximate Bayesian inference," Ph.D. dissertation, Univ. Cambridge, Cambridge, U.K., May 2003.
[134] J. Dauwels, "On variational message passing on factor graphs," in *Proc. IEEE ISIT*, Jun. 2007, pp. 2546–2550.
[135] J. Dauwels, On variational message passing on factor graphs. [Online]. Available: http://www.dauwels.com/localpapers/VMP.pdf
[136] A. V. Zelst and T. C. W. Schenk, "Implementation of a MIMO OFDM-based wireless lan system," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 483–493, Feb. 2004.
[137] W. G. Jeon, K. H. Paik, and Y. S. Cho, "An efficient channel estimation technique for OFDM sytems with transmiter diversity," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun.*, London, U.K., Sep. 2000, vol. 2, pp. 1246–1250.
[138] J. Siew, R. Piechocki, A. Nix, and S. Armour, "A channel estimation method for MIMO-OFDM systems," in *Proc. London Commun. Symp.*, London, U.K., Sep. 2002, pp. 1–4.
[139] M. Shin, H. Lee, and C. Lee, "Enhanced channel estimation technique for MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 261–265, Jan. 2004.
[140] S. Sun, I. Wiemer, C. K. Ho, and T. T. Tjhung, "Training sequence assisted channel estimation for MIMO-OFDM," in *Proc. IEEE Wireless Commun. Netw. Conf.*, New Orleans, LA, USA, Mar. 2003, vol. 1, pp. 38–43.
[141] Y. Gong and K. B. Letaief, "Low complexity channel estimation for space-time coded wideband OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 876–882, Sep. 2003.
[142] Y. Gong and K. B. Letaief, "Low rank channel estimation for space time coded wideband OFDM systems," in *Proc. IEEE VTC*, Atlantic City, NJ, USA, Oct. 2001, pp. 772–776.
[143] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.
[144] C. Suh, C. S. Hwang, and H. Choi, "Preamble design for channel estimation in MIMO-OFDM systems," in *Proc. IEEE GLOBECOM*, San Francisco, CA, USA, Dec. 2003, vol. 1, pp. 317–321.
[145] S. Beyme and C. Leung, "Efficient computation of DFT of Zadoff-Chu sequences," *Electron. Lett.*, vol. 45, no. 9, pp. 461–463, Apr. 2009.
[146] H. Minn and N. A. Dhahir, "Optimal training signals for MIMO-OFDM channel estimation," *IEEE Trans. Wireless Commun.*, vol. 5, no. 5, pp. 1158–1168, May 2006.
[147] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO-OFDM systems in mobile wireless channels," *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp. 1615–1623, Jun. 2003.
[148] T. L. Tung, K. Yao, and R. E. Hudson, "Channel estimation and adaptive power allocation for performance and capacity improvement of multiple-antenna OFDM systems," in *Proc. IEEE Signal Process. Workshop Signal Process. Adv. Wireless Commun.*, Mar. 2001, pp. 82–85.
[149] T. A. Thomas and F. W. Vook, "Broadband MIMO-OFDM channel estimation via near-ML time of arrival estimation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2002, vol. 3, pp. 2569–2572.
[150] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 10, pp. 1553–1560, Oct. 1988.
[151] Z. J. Wang, Z. Han, and K. J. R. Liu, "A MIMO-OFDM channel estimation approach using time of arrivals," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1207–1213, May 2005.
[152] J. Luo, K. R. Pattipati, P. K. Willett, and F. Hasegawa, "Near-optimal multiuser detection in synchronous CDMA using probabilistic data association," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 361–363, Sep. 2001.
[153] M. A. Khojastepour, K. Gomadam, and X. Wang, "Pilot assisted channel estimation for MIMO-OFDM systems using theory of sparse signal recovery," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2009, pp. 2693–2696.
[154] F. Wan, W. P. Zhu, and M. N. S. Swamy, "Semiblind sparse channel estimation for MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 60, no. 6, pp. 2569–2581, Jul. 2011.
[155] F. Gao and A. Nallanathan, "Blind channel estimation for MIMO-OFDM systems via nonredundant linear precoding," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 784–789, Feb. 2007.

[156] C. Shin, R. W. Heath, and E. J. Powers, "Blind channel estimation for MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 670–685, Mar. 2007.

[157] H. Artes, D. Seethaler, and F. Hlawarsch, "Efficient detection algorithms for MIMO channels: A geometrical approach to approximate ML detection," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2808–2820, Nov. 2003.

[158] B. M. Hochwald and S. T. Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.

[159] M. O. Damen, A. Chkeif, and J. C. Belfiore, "Lattice code decoder for space-time codes," *IEEE Commun. Lett.*, vol. 4, no. 5, pp. 161–163, May 2000.

[160] M. O. Damen, H. E. Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2389–2402, Oct. 2003.

[161] Z. Guo and P. Nilsson, "Algorithm and implementation of the K-best sphere decoding for MIMO detection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 491–503, Mar. 2006.

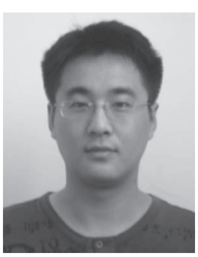
[162] Artist4G, Advanced receiver signal processing techniques: Evaluation and characterization. [Online]. Available: https://ict-artist4g.eu

[163] L. Hanzo, Y. Akhtman, M. Jiang, and L. Wang, *MIMO-OFDM for LTE, WIFI and WIMAX: Coherent Versus Non-Coherent and Cooperative Turbo-Transceivers*. Hoboken, NJ, USA: Wiley, 2010.

[164] Y. Shen and E. F. Martinez, "Channel estimation on OFDM systems," Freescale Semiconductor Inc., Austin, TX, USA, Appl. Notes, AN3059, Jan. 2006.

[165] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.

[166] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599–618, Feb. 2001.

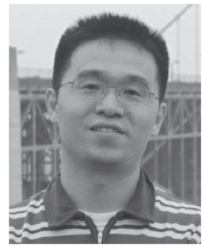


Yinsheng Liu received the B.S. degree from North China Electric Power University, Baoding, China, and the M.S. degree from Beijing Jiaotong University, Beijing, China, in 2007 and 2009, respectively, all in communication and information system. He is currently working toward the Ph.D. degree at Beijing Jiaotong University, Beijing, China.

His research interests include wireless communication system and digital signal processing.

Zhenhui Tan was born on February 18, 1944, in Yangzhou, China. He graduated from Beijing Jiaotong University (BJTU), Beijing, China, and majored in wireless communications in 1967. He received the M.S. degree from BJTU in 1981 and the Ph.D. degree from Southeast University, Nanjing, China, in 1987.

From 1998 to 2008, he was the President of BJTU. From 1995 to 1998, he was the Vice President of BJTU. From 1993 to 1995, he was the Director of the Department of Communications and Signal Control Engineering, BJTU. From 1990 to 1993, he studied as a Visiting Scholar with Mons Technology College, Belgium, and with the University of Waterloo, Waterloo, ON, Canada. He is currently with the Institute of Broadband Wireless Mobile Communications and State Key Laboratory of Rail Traffic Control and Safety, BJTU, China. He is a member of the National Hi-Tech Project's Specialist Committee for three consecutive tenures, a member of the Evaluation Committee for Academic Degree Affairs of the State Council for two consecutive tenures, a member of the Specialist Committee for Postdoctoral Degree Affairs, the Director of the electrical information science and engineering teaching guidance committee of Ministry of Education. He was awarded by State Excellent Coming-Back From-Abroad Scholars, Experts of Outstanding Achievements from the State Council. His research interests include communications and information systems, particularly digital mobile communications, spread spectrum communications, and adaptive filtering algorithms.



Hongjie Hu received the B.Eng. degree from Northwestern Polytechnic University, Xi'an, China, and the M. Eng. degree from the National University of Singapore, Singapore, both in electrical engineering.

He has been with Huawei Technologies for ten years, and works on air interface algorithms design and radio network performance analysis for various types of cellular base stations

Leonard J. Cimini, Jr. (S'77–M'82–SM'89–F'00) received the Ph.D. degree from the University of Pennsylvania, Philadelphia, PA, USA, in 1982.

He worked at AT&T, first in Bell Labs and then AT&T Labs, for 20 years, and has been a Professor with the University of Delaware, Newark, DE, USA, since 2002. He has held several positions within the IEEE Communications Society. He is currently the Director of Journals and, among other publications-related positions, is the founding Editor-in-Chief of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (J-SAC): WIRELESS COMMUNICATIONS SERIES. He was elected an IEEE Fellow in 2000 for contributions to the theory and practice of high-speed wireless communications. For this pioneering work on OFDM for wireless systems, he was given the 2007 James R. Evans Avant Garde Award from the IEEE Vehicular Technology Society and the 2010 Innovators Award from the NJ Inventors Hall of Fame. He has received several other awards from the IEEE Communications Society, including the 2010 Stephen O. Rice Prize, the 2010 Donald W. McLellan Meritorious Service Award, the 2010 Recognition Award from the Wireless Communications Technical Committee, and the 2011 Service Award from the Communication Theory Technical Committee.

Geoffrey Ye Li (S'93–M'95–SM'97–F'06) received the B.S.E. and M.S.E. degrees from Nanjing Institute of Technology, Nanjing, China, in 1983 and 1986, respectively, and the Ph.D. degree from Auburn University, Auburn, AL, USA, in 1994.

He was a Teaching Assistant and then a Lecturer with Southeast University, Nanjing, China, from 1986 to 1991, a Research and Teaching Assistant with Auburn University, Auburn, AL, USA, from 1991 to 1994, and a Postdoctoral Research Associate with the University of Maryland at College Park, MD, USA, from 1994 to 1996. He was with AT&T Labs—Research at Red Bank, NJ, USA, as a Senior and then a Principal Technical Staff Member from 1996 to 2000. Since 2000, he has been with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA, as an Associate Professor and then as a Full Professor. He is also holding the Cheung Kong Scholar title at the University of Electronic Science and Technology of China, Chengdu, China, since March 2006. His general research interests include statistical signal processing and communications, with emphasis on cross-layer optimization for spectral- and energy-efficient networks, cognitive radios and opportunistic spectrum access, and practical issues in LTE systems. In these areas, he has published over 300 refereed journal and conference papers in addition to 25 granted patents. His publications have been cited over 16 000 times from Google Citations and he is listed as a highly cited researcher by Thomson Reuters. He has been involved in editorial activities in more than ten technical journals for the IEEE Communications and Signal Processing Societies. He organized and chaired many international conferences. He has served as Technical Program Vice-Chair of IEEE ICC'03, the Technical Program Cochair of IEEE SPAWC'11, and the General Chair of IEEE GlobalSIP'14. He has been awarded IEEE Fellow status for his contributions to signal processing for wireless communications since 2006. He won the Stephen O. Rice Prize Paper Award in 2010 and the WTC Wireless Recognition Award in 2013 from the IEEE Communications Society. He also received the James Evans Avant Garde Award in 2013 from the IEEE Vehicular Technology Society.