Signal Processing Advances in Communications

Volume II: Trends in Single- and Multi-User Systems

Edited by

P. Stoica, G. Giannakis, Y. Hua, and L. Tong

CONTENTS

Chapter 4

TIME-VARYING FADING CHANNELS [1]

S. BARBAROSSA[1], A. SCAGLIONE[2]

[1] INFOCOM DEPT., UNIVERSITY OF ROME "LA SAPIENZA" - ITALY
[2] DEPT. OF ELECTRICAL ENGINEERING, UNIVERSITY OF MINNESOTA - U.S.A.

sergio@infocom.uniroma1.it, anna@mail.ece.umn.edu

The increasing demand of multimedia communications among mobile users is one of the main motivations for the study of time-varying channels. Examples of mobile communications include cellular communications among users located in cars, trains or airplanes, and digital broadcasting from low Earth orbit (LEO) satellites. Of course, relative motion between transmitter and receiver is one of the most common sources of channel variability, but it is not the only one. Channel fluctuations are also induced by instabilities of the transmit/receive equipments, such as oscillators' phase noise, frequency drifts and sampling jitter, or by the motion of the scatterers composing the transmission medium (especially evident in underwater acoustic links or in radio links involving refraction from the ionosphere, for example). The Doppler effect, which is one of the main causes of channel variability, is directly proportional to both carrier frequency and relative velocity between transmitter and receiver. Hence, both current trends towards higher carrier frequencies and mobility induce higher Doppler shifts and thus faster channel fluctuations. Nonetheless, channel variations can be classified as slow or fast only with respect

[1]THIS RESEARCH WAS SUPPORTED BY NSF WIRELESS INITIATIVE GRANT NO. 99-79443 AND BY US ARMY RESEARCH LABORATORY (ARL) GRANT DAAL01-98-P-0966.

to the transmission rate. Considering that bit rates already in use or foreseen for near future services range from a few hundreds of Kbps up to a few hundreds of Mbps (see, e.g., the Universal Mobile Telecommunication System where the chip rate is 2048 Kbps or the future Mobile Broadband Systems targeting data rates up to 155 Mbps), the channel fluctuations are usually much slower than the transmission rates. However, the situation is more complicated because of coding, spreading or scrambling, which are inevitably required to provide an efficient and reliable use of the transmission resources. In fact, several current systems use spreading codes to combat channel dispersiveness, as in orthogonal frequency division multiplexing (OFDM) for example, or for allowing multiple transmissions over a shared medium, as in cellular systems adopting code division multiple access (CDMA). For example, in the European standard for the Digital Terrestrial Television Broadcasting (DTTB) system that adopts OFDM, the coded symbol has a duration that may be $2,048$ or $8,192$ times greater than the duration of the input information symbol (see, e.g., the so called $2K$ and $8K$ modes of DTTB [27]). Furthermore, in cellular networks based on CDMA, each user has a specific spreading code and each base station has a unique scrambling code. Both codes are composed of several chips. Moreover, in single-rate systems, the number of chips of the spreading code is equal to the maximum number of active users per cell, but in variable-rate services the system can be fully loaded also with a number of users much smaller than the number of chips. Hence, the need of accommodating more and more users within each cell leads inevitably to an increase of the code lengths. As a consequence of spreading and scrambling, the duration of the coded symbol can thus be much greater than the information bit duration. Therefore, even though a channel is slow with respect to the transmission rate, it may be fast with respect to information symbol rate. Under such circumstances, since recovering the information symbols is the ultimate goal of any communication system, the channel fluctuations must be properly taken into account.

In the design of a communication link through a time-varying channel, one may see the channel fluctuations as a source of impairment which must be properly counteracted. However, the most interesting strategy, whenever applicable, consists in converting the channel variability into a useful source of diversity. This approach mirrors the same paradigm shift already observed in the development from narrowband to wideband communications through frequency-selective linear time-invariant (LTI) channels. Narrowband LTI channels are modeled as a multiplicative coefficient and as such they do not introduce any intersymbol interference (ISI). But, if that coefficient is very low the received signal is highly attenuated and the only possibility to counteract this shortcoming is to resort to space diversity. Conversely, wideband links over channels affected by multipath propagation are prone to ISI. However, if the transmission bandwidth is large enough to resolve a certain number of paths, the linear combination of the resolved path yields an SNR gain, without requiring any space diversity. This is indeed the basic idea underlying spread-spectrum systems, which use a bandwidth B much greater than the information symbol rate, and then employ a rake receiver to combine the resolved paths

[71]. In this way, the channel time dispersiveness is turned into a useful source of diversity, using a simple receiver scheme.

Extending the same arguments to linear time-varying (LTV) channels exhibiting dispersiveness possibly along both time and frequency axes, we can exploit advantageously the time and frequency diversity of the channel. This requires i) transmission of signals with sufficient time-bandwidth product to allow us to resolve a sufficient number of paths in both time and frequency domains; ii) accurate estimate of the channel fluctuations; iii) coherent combination of the resolved paths to increase the SNR using, for example, a 2-D RAKE filter [77].

An even more interesting situation arises when the transmitter is able to predict the evolution of the channel on the basis of the past channel behavior, at least within a finite interval. This is possible whenever the transmitter can acquire channel status information (CSI) through feedback channels or exploiting channel reciprocity in time-division duplexing (TDD) links. In the first case, ad hoc low bit rate channels are reserved to exchange system informations such as power, rate, link properties and so on; in the second case, we exploit the simple property that the channel between two users is necessarily the same, irrespective of which user is transmitting or receiving. In both cases it is possible to precode the signal to be transmitted in order to optimize the use of the available resources and thus exploit the inherent channel diversity. Optimal coding strategies, maximizing the information rate for transmissions over flat fading time-varying channels were investigated in [37], [42], [15], for example, whereas time *and* frequency selective channels were analyzed in [64], [63], [9].

In this chapter, far from providing a tutorial presentation of the transmission over time-varying channels which would require much more space, we wish to highlight a few aspects of the problem with the hope of stimulating further developments in this challenging field. Two peculiar aspects of the approach followed in this presentation are: i) analytic (approximate) modeling of the eigenfunctions of underspread LTV channels, and ii) parametric modeling of the channel impulse response. The analytic modeling, even though approximate, has been fundamental to provide physical insight into the optimal coding strategies and to devise simple sub-optimal coding structures. The parametric modeling has been instrumental in quantifying the channel estimation and prediction errors.

The chapter is organized as follows. In Section 4.1 we introduce the channel model describing how different perspectives, i.e. deterministic versus stochastic or parametric versus non-parametric modeling, can all coexist as they shed light on different, yet equally important, aspects of the same problem. A special attention is devoted to modeling the eigenfunctions of underspread channels, as they provide a good model for wireless communication links. In Section 4.2 we derive the optimal linear precoding strategies according to different criteria, such as maximum information rate or minimum mean square error between transmitted and estimated symbols, subject to zero-forcing or average transmit power constraints. Capitalizing on the analytic model of the underspread channel's eigenfunctions, we derive closed form expressions that show how power and bits are distributed as a

function of time and frequency, according to different optimization criteria. The analytic model is also instrumental to devise sub-optimal ways to implement the coding strategy using, for example, the *adaptive* OFDM system, described in Section 4.2.3, that allocates power and bits jointly as a function of both frequency bins and blocks. The channel is initially assumed to be perfectly known. Then we assume, more realistically, that only partial informations are available and we show how it is possible to incorporate the uncertainties about our knowledge on the channel status into the precoding strategy. In particular, in Section 4.2.4 we evaluate the effects of channel prediction errors on the optimal power/bit allocation. Channel estimation and prediction are then analyzed in Section 4.3 where we provide the lower bound on the accuracy of the channel parameters estimates and show how the estimation accuracy depends on number of samples, SNR and channel order estimation.

4.1 Channel model

The most general form for the input/output relationship of a continuous-time (CT) LTV channel with additive noise is [10], [44]:

$$y(t) = \int_{-\infty}^{\infty} h(t,\tau)x(t-\tau)d\tau + v(t) \tag{4.1.1}$$

where $x(t)$ and $y(t)$ are the channel input and output functions, $h(t,\tau)$ is the channel impulse response and $v(t)$ is the noise. In general, given the high complexity of the physical interactions characterizing the transmission through a real channel (e.g. reflections, scattering, refraction or diffractions), the most appropriate modeling of the impulse response is probabilistic, so that $h(t,\tau)$ is a 2D random process. However, in several cases of practical interest, the realizations of this random process can be described very accurately by a parametric model whose parameters represent physically meaningful quantities, such as delays, Doppler frequencies and reflection coefficients. Therefore, both deterministic and stochastic approaches are equally useful in describing a time-varying channel, even though they embrace different aspects: The stochastic model is better suited for describing global behaviors, whereas the deterministic one is more useful to study the transmission through a specific channel realization. Modeling the parameters of a multipath channel as random variables, for example, provides a simple yet important random channel model. Parametric modeling is especially important in devising channel estimation and tracking algorithms (see e.g. [34]). For all these reasons, in the next section we start with a deterministic characterization whereas the ensuing section will be devoted to some of the basic concepts of random modeling.

4.1.1 Deterministic models

We recall here the basic relationships of continuous-time and equivalent discrete-time base-band channels.

Continuous-time model

Following the same notation introduced in the pioneering work of Bello [10], any LTV channel can be fully characterized by its impulse response $h(t,\tau)$, introduced in (4.1.1), or by any of the following system functions:

i) time-varying transfer function

$$H(t,f) := \int_{-\infty}^{\infty} h(t,\tau) e^{-j2\pi f\tau} d\tau; \tag{4.1.2}$$

ii) delay-Doppler spread function

$$S(\nu,\tau) := \int_{-\infty}^{\infty} h(t,\tau) e^{-j2\pi\nu t} dt; \tag{4.1.3}$$

iii) output Doppler-spread function

$$Q(\nu,f) := \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(t,\tau) e^{-j2\pi(\nu t+f\tau)} dt d\tau. \tag{4.1.4}$$

Unless otherwise stated, in the following we will denote the function $S(\nu,\tau)$ simply as the channel spread function. The function $Q(\nu,f)$ is useful to derive the dual input/output relationship (4.1.1) in the frequency domain:

$$Y(f) = \int_{-\infty}^{\infty} Q(f-\nu,\nu) X(\nu) d\nu = \int_{-\infty}^{\infty} Q(\nu,f-\nu) X(f-\nu) d\nu, \tag{4.1.5}$$

where $X(f)$ and $Y(f)$ denote the Fourier transform (FT) of the input and output signals, respectively. The time-varying transfer function $H(t,f)$ is related to the spread function $S(\nu,\tau)$ by a two-dimensional Fourier Transform:

$$H(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(\nu,\tau) e^{j2\pi(\nu t-f\tau)} d\tau\, d\nu. \tag{4.1.6}$$

From (4.1.6), it is clear that the variability of $H(t,f)$ can be measured by checking the spreading of $S(\nu,\tau)$, which can be quantified through its moments. In general, since the spread function may be concentrated around a point, in the (ν,τ) plane, different from the origin, it is more meaningful to measure its spread using the central moments. Considering also that $S(\nu,\tau)$ is complex, we measure the spread of a channel through its normalized absolute central moments, defined as

$$m_S^{(k,l)} := \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |S(\nu,\tau)||\tau-t_0|^k |\nu-f_0|^l d\nu d\tau}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |S(\nu,\tau)| d\nu d\tau} \tag{4.1.7}$$

The values t_0 and f_0 can be chosen to minimize, for example, $m_S^{(1,0)}$ and $m_S^{(0,1)}$, separately. In such a case, t_0 is equal to the median value of $S(\tau) := \int_{-\infty}^{\infty} |S(\nu,\tau)| d\nu$, and f_0 is the median value of $S(\nu) := \int_{-\infty}^{\infty} |S(\nu,\tau)| d\tau$.

In general, for any given channel input $x(t)$, a shift of the spread function causes a corresponding shift, in time and frequency, of the corresponding channel output. More precisely, the following property holds true.

P0: Denoting by $y(t)$ the output of a channel having a spread function $S(\nu,\tau)$ corresponding to the input $x(t)$, the output $\bar{y}(t)$ of the channel having spread function $\bar{S}(\nu,\tau) = S(\nu - f_0, \tau - t_0)e^{-j2\pi(\nu-f_0)t_0}$, corresponding to the same input $x(t)$, is a time and frequency shifted replica of $y(t)$, i.e. $\bar{y}(t) = y(t-t_0)e^{j2\pi f_0 t}$.
Proof: Combining (4.1.1) and (4.1.3) we have

$$y(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(\nu,\tau)x(t-\tau)e^{j2\pi\nu t}d\nu d\tau, \tag{4.1.8}$$

and, similarly

$$\begin{aligned}\bar{y}(t) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(\nu-f_0,\tau-t_0)e^{-j2\pi(\nu-f_0)t_0}x(t-\tau)e^{j2\pi\nu t}d\nu d\tau \qquad (4.1.9)\\ &= e^{j2\pi f_0 t}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(\xi,\theta)x(t-t_0-\theta)e^{j2\pi\xi(t-t_0)}d\xi d\theta = e^{j2\pi f_0 t}y(t-t_0).\end{aligned}$$

Therefore, we may concentrate, without any loss of generality, on channels whose spread function is centered around the origin of the (ν,τ) plane.

We say that a channel is *underspread* if

$$m_S^{(k,0)}m_S^{(0,l)} \ll 1, \forall k,l \geq 1. \tag{4.1.10}$$

This means that the spread function of an underspread channel must be concentrated at least along one axis, either along ν or along τ, or along both (among these situations, the case where the spread function is concentrated along both axes is probably the least interesting because it corresponds to an almost flat fading channel).

In wideband mobile communications, the channel can be modeled as the superposition of a discrete number of paths [44]. When the transmitted signal has a bandwidth B greater than the channel coherence bandwidth B_c (see, e.g. (4.1.36) for its definition), a certain number, let us say K, of paths can be resolved in time, thus leading to the following model

$$h(t,\tau) = \sum_{k=0}^{K-1} h_k(t)\delta(\tau - \tau_k(t)), \tag{4.1.11}$$

where $h_k(t)$ and $\tau_k(t)$ denote the temporal variations of the complex amplitude and delay associated to the kth path, respectively. In most cases, the delays do not vary significantly over a symbol period so that the time variations of $h_k(t)$ can be fitted by a linear combination of sinusoids [47], [19], [34] [2]:

$$h_k(t) = \sum_{p=0}^{P-1} h_{k,p}e^{j2\pi f_{k,p}t}. \tag{4.1.12}$$

[2]The model proposed in [34] is slightly different from (4.1.12) because it assumes that the expansion basis composed of the complex exponentials is the same for all k's. This assumption was instrumental in [34] to devise blind channel estimation and equalization techniques.

Using the constant delay assumption and combining (4.1.12) with (4.1.11), the impulse response can also be written as

$$h(t,\tau) = \sum_{q=0}^{Q-1} h_q e^{j2\pi f_q t} \delta(\tau - \tau_q) \tag{4.1.13}$$

where $Q = PK$. This expression models the channel as a superposition of Q paths, each one characterized by the triplet (h_q, τ_q, f_q). We assume that there are no paths having the same pair of delay/Doppler values (τ_q, f_q) (otherwise the two paths are grouped to form a single one). However, there may be different paths having the same delay or the same Doppler shift.

The spread function and time-varying transfer functions corresponding to (4.1.13) are:

$$S(\nu,\tau) = \sum_{q=0}^{Q-1} h_q \delta(\tau - \tau_q)\delta(\nu - f_q) \tag{4.1.14}$$

and

$$H(t,f) = \sum_{q=0}^{Q-1} h_q e^{j2\pi(f_q t - f\tau_q)}, \tag{4.1.15}$$

respectively. Since practical values of both τ_q and f_q are bounded, (4.1.14) reveals the finite support nature of the channel spread function in both Doppler and delay domains. Using (4.1.7), the normalized absolute moments of a spreading channel are (we assume for simplicity of notation, but without loss of generality, according to ***P0***, that the median values t_0 and f_0 are equal to zero):

$$m_S^{(k,l)} = \frac{\sum_{q=0}^{Q-1} |h_q||\tau_q|^k |\nu_q|^l}{\sum_{q=0}^{Q-1} |h_q|}. \tag{4.1.16}$$

We say that a multipath channel is underspread if

$$\max_{p,q}\{|\tau_p \nu_q|\} \ll 1 \tag{4.1.17}$$

or if

$$m_S^{(k,0)} m_S^{(0,l)} \ll 1, \forall k,l \geq 1. \tag{4.1.18}$$

The two definitions are not exactly equivalent, because only the second one takes the path amplitudes into account.

Discrete-time model

In digital communication systems using linear modulation, the baseband transmitted signal can be expressed as

$$x(t) = \sum_{k=-\infty}^{\infty} x[k] g_T(t - kT_s), \tag{4.1.19}$$

where $x[k]$ is the (generally complex) kth transmitted symbol and $g_T(t)$ is the transmit lowpass filter whose bandwidth is directly proportional to the symbol rate

$1/T_s$. The impulse response $g_T(t)$ has a Nyquist characteristic and it is usually a root raised cosine filter [71]. The channel output can then be written as

$$z(t) = \sum_{k=-\infty}^{\infty} x[k] \int_{-\infty}^{\infty} h(t,\tau) g_T(t-\tau-kT_s) d\tau + w(t), \tag{4.1.20}$$

where $w(t)$ is additive noise. The received signal is demodulated, low-pass filtered and sampled. Denoting by $g_R(t)$ the impulse response of the receive low-pass filter, the baseband received signal is

$$y(t) = \sum_{k=-\infty}^{\infty} x[k] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_R(t-\theta) g_T(\theta-\tau-kT_s) h(\theta,\tau) d\tau d\theta + v(t), \tag{4.1.21}$$

where $v(t)$ is the filtered noise. Hence, sampling $y(t)$ at symbol period T_s, we obtain the sequence

$$y[n] := y(nT_s) = \sum_{k=-\infty}^{\infty} h[n, n-k] x[k] + v[n], \tag{4.1.22}$$

where we have introduced the equivalent discrete-time impulse response

$$h[n, n-k] := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_R(nT_s-\theta) g_T(\theta-\tau-kT_s) h(\theta,\tau) d\tau d\theta. \tag{4.1.23}$$

Equation (4.1.22) is the discrete-time counterpart of (4.1.1).

To gain better insight into the transmission through LTV channels, it is useful to express (4.1.23) in the frequency domain. Specifically, introducing the transfer functions $G_T(f)$ and $G_R(f)$ of the transmit and receive filters and using the output Doppler-spread function (4.1.4), we may rewrite (4.1.23) as

$$h[n, n-k] := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G_R(f) G_T(\nu) Q(f-\nu, \nu) e^{j2\pi(nf-\nu k)T_s} d\nu df. \tag{4.1.24}$$

Using now the multipath channel model (4.1.13), we get

$$h[n, n-k] = \sum_{q=0}^{L} h_q e^{j2\pi f_q n T_s} \int_{-\infty}^{\infty} G_R(\nu+f_q) G_T(\nu) e^{j2\pi\nu((n-k)T_s+\tau_q)} d\nu. \tag{4.1.25}$$

Substituting the transmit and receive transfer functions $G_T(f)$ and $G_R(f)$ in (4.1.25), we obtain the equivalent discrete-time impulse response in the most general case. To comprehend some of the basic features of the discrete-time equivalent channels, it is useful to analyze the simple case where both transmit and receive shaping filters are ideal lowpass filters. In particular, setting $G_T(f) = G_R(f) = \sqrt{T_s} rect(fT_s)$, where the rectangular function $rect(f)$ is equal to one for $|f| < 1/2$ and is null otherwise, (4.1.25) gives rise to the following DT impulse response

$$h[n,k] = \sum_{q=0}^{L} h_q [1-|f_q|T_s]^+ e^{j\pi\nu_q(2n-k+\theta_q)} sinc[\pi(1-|\nu_q|)(k-\theta_q)], \tag{4.1.26}$$

where $[x]^+ \equiv \max(x, 0)$ and we have introduced the normalized delay $\theta_q := \tau/T_s$ and Doppler $\nu_q := f_q T_s$. Since $\max_q |\nu_q| < 1$, it is clear from (4.1.26) that the components corresponding to higher Doppler shifts ν_q are more attenuated. This happens because part of their energy falls outside of the receive filter bandwidth. In fact, since the transit through an LTV channel increases the bandwidth of the signal, the receive filter should have a bandwidth greater than $1/T_s$ to keep all the useful energy and, consequently, the sampling rate should also be higher than $1/T_s$, to avoid any loss of information. However, in practice $f_q T_s \ll 1$, so that usually the receive filter bandwidth and sampling rate can be maintained equal to $1/T_s$, without any appreciable loss. As a simple numerical example, using a carrier frequency of 2 GHz in a link of 1 Mbps between terminals in relative motion at a velocity of $v = 150$ Km/hr, the maximum normalized Doppler shift is approximately $2.8 \cdot 10^{-4}$. This explains why in most practical systems the receiver bandwidth and sampling rates are not higher than $1/T_s$, for the gain obtainable otherwise is not worth of the extra complications related to re-sampling. Therefore, (4.1.26) can be approximated with negligible error as

$$h[n, k] \simeq \sum_{q=0}^{L} h_q e^{j\pi\nu_q(2n-k+\theta_q)} sinc[\pi(k - \theta_q)]. \tag{4.1.27}$$

A convenient form for expressing the I/O relationship in the DT case is the matrix form. Specifically, in the transmission over an FIR channel of order L, each block of P received symbols $\boldsymbol{y}(n) := (y(nP), \ldots, y(nP + P - 1)^T)$ is linearly related to the corresponding block of R transmitted symbols $\boldsymbol{x}(n) := (x(nR), \ldots, x(nR+R-1))^T$ through the matrix relationship

$$\boldsymbol{y}(n) = \boldsymbol{H}(n)\boldsymbol{x}(n) + \boldsymbol{v}(n), \tag{4.1.28}$$

where $Q = P+L$, $\boldsymbol{H}(n)$ is the $P \times R$ channel matrix whose (i, j) entry is $\{\boldsymbol{H}(n)\}_{i,j} = h[nP + i, i - j]$ and $\boldsymbol{v}(n) := (v(nP), \ldots, v(nP + P - 1)^T)$ is additive noise.

Usually the information symbols are precoded before transmission, to make the received data more resilient against channel dispersiveness and allow multi-user communications over a shared medium. In this chapter we will only consider linear encoding. To simplify synchronization and symbols decoding, we also adopt a block encoding and insert guard intervals to avoid inter-block interference (IBI). Specifically, if the DT channel has an impulse response of finite order L, this is achieved by parsing the input sequence into consecutive blocks of length M and padding guard intervals between consecutive blocks of length at least equal to L. At the receiver, discarding the guard intervals we get rid of the IBI. Of course Inter-Symbol Interference (ISI) is still present. However, redundant linear precoding allows for zero-forcing symbol recovery, regardless of the FIR channel parameters [80]. Linear block coding with null guard intervals can be described through the encoding and decoding matrices $\boldsymbol{F}(n)$ and $\boldsymbol{G}(n)$, so that the nth transmitted block is $\boldsymbol{x}(n) = \boldsymbol{F}(n)\boldsymbol{s}(n)$, whereas the nth decoded symbols block is

$$\hat{\boldsymbol{s}}(n) = \boldsymbol{G}(n)\boldsymbol{H}(n)\boldsymbol{F}(n)\boldsymbol{s}(n) + \boldsymbol{G}(n)\boldsymbol{v}(n). \tag{4.1.29}$$

In Section 4.2 we will show how to optimize the choice of both $\boldsymbol{F}(n)$ and $\boldsymbol{G}(n)$, depending on the amount of channel status information (CSI) available at the transmitter side and on the adopted optimization criterion.

4.1.2 Stochastic models

The deterministic models are useful to describe the properties of specific channel realizations, but are too simple to provide a global characterization of practical time-varying channels. Modeling the channel impulse response as a 2D stochastic process gives many more degrees of freedom which can provide a better matching between mathematical models and experimental results. Furthermore, only stochastic modeling allows the computation of global performance parameters such as outage probability or information outage probability, defined respectively as the probability that the bit error rate exceeds a prescribed value and the probability that the mutual information exceeds a certain rate.

In general, the full characterization of a time-varying random channel requires the joint probability density function (pdf) of any set of random variables extracted from the channel impulse response. However, the determination of such a pdf may be a too difficult task, especially when the channel's probabilistic behavior departs from the Gaussian model. A less stringent, but still meaningful approach involves the determination of the correlation of the system functions introduced in the previous section. We recall a few basic definitions here, which will be used later on. The interested reader may refer to [10], [45], [44] or [70] for more details. Introducing the correlation of the impulse response $R_h(t_1, t_2; \tau_1, \tau_2) := E\{h^*(t_1, \tau_1)h(t_2, \tau_2)\}$ and using Bello's system functions and terminology [10], a channel is wide sense stationary (WSS) if

$$R_h(t_1, t_2; \tau_1, \tau_2) = R_h(\Delta t; \tau_1, \tau_2) \tag{4.1.30}$$

with $\Delta t = t_2 - t_1$. This implies that signals arriving with different Doppler shifts are uncorrelated [44]. By duality, a channel is with uncorrelated scattering (US) if signals arriving with different delays are uncorrelated. In formulas

$$R_h(t_1, t_2; \tau_1, \tau_2) = P_h(t_1, t_2; \tau_2)\delta(\tau_2 - \tau_1). \tag{4.1.31}$$

Combining the two properties, a channel is wide sense stationary with uncorrelated scattering (WSSUS) if

$$R_h(t_1, t_2; \tau_1, \tau_2) = P_h(\Delta t; \tau_2)\delta(\tau_2 - \tau_1). \tag{4.1.32}$$

A practical wireless channel cannot be assumed to be stationary. However, the stationary model is well suited for a small scale characterization, i.e. for short periods of time or, equivalently, for motions over a small geographical area. The large scale behavior can then be captured by examining the small scale statistics across consecutive time intervals (geographical areas). This class of channels has been denoted as quasi- WSSUS (QWSSUS) [10] and is a fairly good model for practical channels.

From a practical point of view, it is better to quantify the main channel characteristics with a few numbers, rather than with functions. Two particularly important parameters are the channel coherence time and bandwidth. Dealing with

WSSUS channels and introducing the function $P_h(\tau) := P_h(0;\tau)$, the channel *average delay* and *delay spread* are defined as

$$D := \frac{\int_0^\infty \tau P_h(\tau)d\tau}{\int_0^\infty P_h(\tau)d\tau}, \qquad \sigma_t := \sqrt{\frac{\int_0^\infty (\tau - D)^2 P_h(\tau)d\tau}{\int_0^\infty P_h(\tau)d\tau}}, \tag{4.1.33}$$

respectively. In [45] it was proved that the correlation coefficient between signals separated by Δf Hz and Δt sec is

$$\rho(\Delta f, \Delta\tau) = \frac{J_0^2(2\pi f_m \Delta t)}{1 + (2\pi\Delta f \sigma_t)^2} \tag{4.1.34}$$

where $J_0(x)$ is the zero order Bessel function of the first kind and $f_m = v/\lambda$ is the maximum Doppler frequency (v is the relative transmit/receive velocity along their line of sight and λ is the transmit wavelength). The channel coherence time and bandwidth quantify the spread of $\rho(\Delta f, \Delta\tau)$ around the origin. Specifically, the *coherence time* is

$$T_c = \frac{9}{16\pi f_m} \tag{4.1.35}$$

and the *coherence bandwidth* is

$$B_c = \frac{1}{2\pi\sigma_t}. \tag{4.1.36}$$

A channel is said to be *frequency-selective* if the transmission bandwidth is greater than the coherence bandwidth and it is *time-selective* if the coded symbol duration exceeds the channel coherence time.

Over a relatively small scale, the amplitude of the dominant paths can be often characterized statistically by a Gaussian pdf, by virtue of the central limit theorem. In fact, given a finite receiver bandwidth B and symbol duration T_s, each dominant path having complex amplitude h_q, delay τ_q and Doppler shift f_q, is equal to the sum of all the components having delays that differs from τ_q by an amount smaller than $1/2B$ and a Doppler differing from f_q by less than $1/2T_s$, in absolute sense. If the number of such components is sufficiently high, we may invoke the central limit theorem to model h_q as a complex Gaussian random variable. Depending on whether or not there are line-of sight paths, the modulus of each h_q is then a Rice or a Rayleigh random variable. However, especially in the presence of a few dominant scatterers or of shadowing effects, the Gaussian model becomes less appropriate and the pdf of the amplitudes of each path component is better fitted by Nakagami, Weibull or log-normal laws. In an effort to conjugate small and large scale properties to better characterize a QWSSUS channel, an appropriate model for the amplitude of the dominant reflections is the product model $z(t) = x(t)y(t)$, where $x(t)$ is the fast fading envelope, whereas $y(t)$ is the slow fading envelope. An example of such a model is the Suzuki's distribution where $x(t)$ is Rayleigh and $y(t)$ is log-normal [89]. An interesting further development could be based on the use of complex correlated non-Gaussian processes, such as the spherically invariant random processes [99] or the complex correlated Weibull process [30], for example.

Before concluding this section, it is worth recalling that (4.1.13) can also provide a simple, yet effective, model for the simulation of a real time-varying channel.

In fact, if measurements of the spread function are available, we can generate an LTV channel fitting the real data in the following way: Starting from a certain number of paths Q, we generate Q pairs as random variables (τ_q, f_q) whose joint pdf is proportional to the modulus of the measured spread function [97]. Each pair is then associated to a complex coefficient h_q whose modulus is Rayleigh, Rice or Nakagami, and whose phase is uniform in $[-\pi, \pi)$. This simulation model, while preserving the parametric structure, can provide a fairly good fit with the experimental measurements. Another useful approach for the simulation of LTV channels consists in modeling the time-varying amplitudes $h_k(t)$ in (4.1.11) as autoregressive (AR) processes [91], [92]. Both superposition of complex exponentials or AR model provide a finite parameterization of the channel impulse response that can be used not only for channel simulation but, more interestingly, also for channel estimation and equalization [36], [34].

4.1.3 Channel singular functions

Since in general the operator describing an LTV channel is not self-adjoint, we cannot define the eigenfunctions of an LTV channel but we can guarantee the existence of right (input) and left (output) channel singular functions under proper conditions on the impulse response. The singular functions are the continuous-time counterparts of the left and right singular vectors resulting from the singular value decomposition (SVD) of the channel matrix. In the following we will use Gallager's notation [33], slightly modified to make it coherent with (4.1.1). In particular, we will assume as in [33] that the channel impulse response is square integrable. In this case, it is known [33] that there exists a sequence of non negative real numbers λ_i and two sets of orthonormal functions $u_\lambda(t)$ and $v_\lambda(t)$ such that the following pair of integral equations hold true [3]:

$$\lambda u_\lambda(t) = \int_{-\infty}^{\infty} h(t, t-\tau) v_\lambda(\tau) d\tau \tag{4.1.37}$$

and

$$\lambda v_\lambda(\tau) = \int_{-\infty}^{\infty} h^*(t, t-\tau) u_\lambda(t) dt. \tag{4.1.38}$$

Inserting (4.1.37) in (4.1.38), we also get

$$\lambda^2 v_\lambda(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h^*(t, t-\tau) h(t, t-\theta) v_\lambda(\theta) d\theta dt \tag{4.1.39}$$

so that $v_\lambda(\tau)$ is the eigenfunction of the system whose kernel is

$$\tilde{h}(\tau, \theta) := \int_{-\infty}^{\infty} h^*(t, t-\tau) h(t, t-\theta) dt. \tag{4.1.40}$$

Similarly, substituting (4.1.38) in (4.1.37), $u_\lambda(t)$ is an eigenfunction of the system with kernel

$$\bar{h}(t, \theta) := \int_{-\infty}^{\infty} h(t, t-\tau) h^*(\theta, \theta-\tau) d\tau. \tag{4.1.41}$$

[3] We can assume without any loss of generality that λ is real and non-negative.

In practice, there are at least two quite common situations where $h(t,\tau)$ is not square-integrable: i) LTI channels, where $h(t,\tau)$ is constant along t; and ii) multipath channels with specular reflections, where $h(t,\tau)$ contains Dirac pulses. Both situations are considered as limiting cases, as in [33].

The discrete time counterpart of (4.1.37) and (4.1.38) is the SVD that allows us to factorize the channel matrix $\boldsymbol{H}(n)$ in (4.1.28) as

$$\boldsymbol{H}(n) = \boldsymbol{U}(n)\boldsymbol{\Lambda}(n)\boldsymbol{V}^H(n), \tag{4.1.42}$$

where the columns of $\boldsymbol{U}(n)$ and $\boldsymbol{V}(n)$ are the left and right channel singular vectors associated to the singular values contained in the diagonal matrix $\Lambda(n)$, relative to the channel status corresponding to the nth transmitted block. Dealing with LTV channels, we find useful to extend the concept of system eigenfunction as follows. It is well known that the eigenfunctions of LTI systems are complex exponentials $\exp(j2\pi f_0 t)$ that remain unaltered after passing through the channel, except for a change of amplitude and for a delay. Generalizing this concept to the LTV case, we name generalized eigenfunctions (or G-eigenfunctions) those input signals $w_\lambda(t)$ whose corresponding output is proportional, through the corresponding eigenvalue, to a shifted replica in time *and* frequency of the input. In formulas, a G-eigenfunction $w_\lambda(t)$, if it exists, must satisfy the following identity:

$$\lambda w_\lambda(t-t_0)e^{j\psi_0}e^{j2\pi f_0 t} = \int_{-\infty}^{\infty} h(t,\tau)w_\lambda(t-\tau)d\tau. \tag{4.1.43}$$

We are now able to state the following properties valid in approximate form for *underspread* channels, with an approximation error directly proportional to the moments of the channel spread function and inversely proportional to the corresponding eigenvalue (see App.4.5.1 for the details).

***P1**: The right singular function $v_\lambda(t)$ corresponding to the singular value λ of an LTV channel with small frequency spread, i.e. $\max_q |\nu_q| \ll 1$, can be approximated by the multicomponent signal*

$$v_\lambda(t) = \begin{cases} \sum_{k=1}^{K_\lambda} A_k^{(\lambda)} e^{j\phi_k^{(\lambda)}(t)}, & t \in I_k^{(\lambda)} \\ 0 & \textit{elsewhere} \end{cases} \tag{4.1.44}$$

whose components have a constant amplitude, within a certain time interval, and an instantaneous frequency $f_k^{(\lambda)}(t) = \dot{\phi}_k^{(\lambda)}(t)/2\pi$ resulting from the solution of the following equation

$$\left| H\left(t, f_k^{(\lambda)}(t)\right)\right|^2 = \lambda^2, \tag{4.1.45}$$

where $H(t,f)$ is the channel time-varying transfer function, $\dot{\phi}_k^{(\lambda)}(t)$ is the time derivative of $\phi_k^{(\lambda)}(t)$, $I_k^{(\lambda)}$ denotes the support over t where (4.1.45) admits the k-th real solution, K_λ denotes the number of solutions of (4.1.45) and the amplitudes $A_k^{(\lambda)}$ are chosen to enforce the unit norm of the singular functions. Similarly, the right singular function $v_\lambda(t)$ corresponding to the singular value λ of an LTV channel with small delay spread, i.e. with $\max_q |\tau_q| \ll 1$, can be approximated by the

multicomponent signal whose spectrum is

$$V_\lambda(f) = \begin{cases} \sum_{k=1}^{K_\lambda} C_k^{(\lambda)} e^{j\Phi_k^{(\lambda)}(f)}, & f \in I_k^{(\lambda)} \\ 0 & elsewhere \end{cases} \tag{4.1.46}$$

whose components have a spectrum with a constant amplitude $|C_k|$*, within a certain frequency interval, and a group delay* $t_k^{(\lambda)}(f) = -\dot{\Phi}_k^{(\lambda)}(f)/2\pi$ *resulting from the solution of the following implicit equation*

$$\left| H\left(t_k^{(\lambda)}(f), f\right)\right|^2 = \lambda^2. \tag{4.1.47}$$

P2*: The existence of a real solution for (4.1.45) requires that the modulus of the channel singular values* λ *be bounded between the following limits:*

$$\min_{t,f} |H(t,f)| \leq \lambda \leq \max_{t,f} |H(t,f)|. \tag{4.1.48}$$

P3*: The left singular functions are approximately proportional to a time and frequency shifted version of the right singular functions, i.e.*

$$u_\lambda(t) = e^{j\Psi_0} e^{j2\pi\nu_0 t} v_\lambda(t - \tau_0). \tag{4.1.49}$$

P4*: The channel G-eigenfunctions* $w_\lambda(t)$ *exist and have a Fourier Transform* $W_\lambda(f)$ *which has approximately the same form as* $w_\lambda(t)$*, except for a change of amplitude and scale.*

Remarks: i) adopting our definition of G-eigenfunctions, (4.1.49) states that $v_\lambda(t)$ is a G-eigenfunction of the channel; ii) from (4.1.44) and (4.1.45) it turns out that the functions $v_\lambda(t)$ associated to different eigenvalues have components whose instantaneous frequencies are non-overlapping curves $f = f_\lambda(t)$ of the time-frequency plane (t, f). Since the components of $v_\lambda(t)$ have also a constant amplitude (at least in a certain time interval), the Wigner-Ville Distribution (WVD) of functions $v_\lambda(t)$ (see section tf-analysis for the time-frequency analysis of the channel eigenfunctions)) corresponding to distinct eigenvalues are non-overlapping. As a consequence, using Moyal's formula (4.5.19), the class of functions given by (4.1.44) and (4.1.45) is a class of orthonormal functions.

In the technical literature there are two important classes of contributions to the theoretical analysis of the eigenfunctions of slowly-varying operators: [50], [51], [61] on one side and [85] on the other side. In the first class of works, it was proved that the eigenfunctions of underspread operators can be approximated by well localized functions in the time-frequency plane, with an error upper bounded by the spread function moments; however there was not a direct link between the shape of the eigenfunctions and the time-varying transfer function. In [85] there is an analytic result very close to ***P1***, but valid only for hermitian operators and for slowly-varying operators. Conversely, it is important to remark that ***P1*** is valid also for channels exhibiting fast fluctuations, provided that $|H(t,f)|$ does not vary too fast along *both* time and frequency axes or, equivalently, provided that $S(\nu,\tau)$ is not too spread

in both directions ν and τ. Furthermore, even though we have not addressed this case here, the analytic model described here can be generalized to channels whose spread function is concentrated around a curve in the delay-Doppler domain.

We verify now the validity of properties ***P1÷P4*** analytically for a few simple channels whose eigenfunctions can be derived in closed form and then we will analyze the most general case numerically.

Two-ray multipath channel

Let us consider a multipath channel composed of two rays, i.e. (4.1.13) with $Q = 2$. We assume that the delays τ_k are not equal to each other (this situation will be considered in the example referring to the multiplicative channel). In such a case, the channel G-eigenfunction must be a solution of

$$\lambda e^{j\psi_0} w_\lambda(t - t_d) e^{j2\pi f_d t} = h_0 w_\lambda(t - \tau_0) e^{j2\pi f_0 t} + h_1 w_\lambda(t - \tau_1) e^{j2\pi f_1 t}. \tag{4.1.50}$$

Setting $t_d = \tau_0$ and $f_d = f_0$, we may rewrite (4.1.50) as

$$\lambda w_\lambda(t - \tau_0) = h_0 w_\lambda(t - \tau_0) + h_1 w_\lambda(t - \tau_1) e^{j2\pi(f_1 - f_0)t}. \tag{4.1.51}$$

Setting $\theta = t - \tau_0$, we have

$$\lambda w_\lambda(\theta) = h_0 w_\lambda(\theta) + h_1 w_\lambda(\theta - \Delta\tau) e^{j2\pi\Delta f(\theta + \tau_0)}, \tag{4.1.52}$$

where $\Delta f := f_1 - f_0$ and $\Delta\tau := \tau_1 - \tau_0$. It is straightforward to verify, by direct substitution, that the solution of (4.1.52) is given by the functions

$$w_\lambda(t) = e^{j2\pi \frac{\alpha}{\Delta\tau} t} e^{j\pi \frac{\Delta f}{\Delta\tau} t^2} \tag{4.1.53}$$

parameterized with respect to the variable α, which is related to λ by

$$\lambda e^{j\psi_0} = h_0 + h_1 e^{j2\pi\alpha} e^{-j\pi\Delta f(\tau_1 + \tau_0)}, \tag{4.1.54}$$

where the phase ψ_0 is chosen in order to have a real non-negative value for λ. Therefore, the *G-eigenfunctions of 2-rays channels are chirp signals* whose sweep rate is $\Delta f/\Delta\tau$. Furthermore, the eigenvalues are related to the channel parameters by (4.1.54). We can now verify ***P1÷P3***. In fact, the channel transfer function is

$$H(t, f) = h_0 e^{j2\pi(f_0 t - f\tau_0)} + h_1 e^{j2\pi(f_1 t - f\tau_1)} = e^{j2\pi(f_0 t - f\tau_0)}(h_0 + h_1 e^{j2\pi(\Delta f t - \Delta\tau f)}) \tag{4.1.55}$$

and the contour lines of $|H(t, f)|$ relative to the level λ are described by the equation

$$\Delta\tau\ f - \Delta f\ \tau = \alpha, \tag{4.1.56}$$

where the parameter α is related to λ by

$$\lambda e^{j\varphi_0} = h_0 + h_1 e^{j2\pi\alpha}, \tag{4.1.57}$$

where φ_0 is any constant phase. Solving (4.1.56) for f and using (4.1.45), the instantaneous frequency of the channel eigenfunctions components is

$$f = f_\lambda(t) = \frac{\Delta f}{\Delta\tau} t + \frac{\alpha}{\Delta\tau} \tag{4.1.58}$$

so that, according to (4.1.44), the channel eigenfunctions are

$$v_\lambda(t) = e^{j2\pi \frac{\alpha}{\Delta\tau} t} e^{j\pi \frac{\Delta f}{\Delta\tau} t^2}. \tag{4.1.59}$$

Notice that this expression coincides with (4.1.53). Furthermore, if the channel is underspread, i.e. $(\tau_1+\tau_0)\Delta f \ll 1$, the channel eigenvalue obtained through (4.1.54) or (4.1.57) coincide by simply setting $\varphi_0 = \psi_0$.

We can also verify the validity of ***P4*** in this case, by recalling that the Fourier Transform (FT) of chirp signals has also approximately a chirp-like behavior [69].

LTI channels

The property that the eigenfunctions of LTI systems assume the form $\exp(j2\pi ft)$ can be deduced as a particular case of ***P1***. In fact, the time-varying transfer function of an LTI channel is constant along t and thus (4.1.44) and (4.1.45) predict that the instantaneous frequency of the channel eigenfunctions must be constant.

Multiplicative channel

A multiplicative channel is characterized by the following input-output relationship

$$y(t) = m(t)x(t) \tag{4.1.60}$$

or, incorporating a possible delay, $y(t) = m(t)x(t-t_0)$. The impulse response of a multiplicative channel assumes the form $h(t,\tau) = \delta(\tau-\tau_0)m(t)$ and its transfer function is $H(t,f) = e^{-j2\pi\tau_0 f}m(t)$. As an example, the multipath model (4.1.13) degenerates into a multiplicative channel when the delays are all equal to each other, e.g. $\tau_k = \tau_0$, $\forall k$. In such a case,

$$h(t,\tau) = \delta(\tau-\tau_0)\sum_{q=0}^{Q-1} h_q e^{j2\pi f_q t}. \tag{4.1.61}$$

Since $|H(t,f)|$ is constant along f, according to (4.1.46) and (4.1.47), the group delay is constant and thus the spectrum of the channel eigenfunctions has constant amplitude and linear phase and thus the eigenfunctions are Dirac pulses. In fact, if the input is $x(t) = \delta(t-\tau_0)$, the corresponding output $y(t)$ is proportional to the input and it is $y(t) = m(\tau_0)\delta(t-\tau_0) = m(\tau_0)x(t)$.

Before concluding this section, we wish to remark that even though the LTI and multiplicative channels are characterized by impulse responses which are not square-integrable, the models (4.1.44) and (4.1.46) in these two cases hold exactly.

4.1.4 Time-frequency analysis of LTV channels' eigenfunctions

In general, besides the few examples considered so far, it is not easy to derive the eigenfunctions of a general LTV channel analytically. Given the central role played by the instantaneous frequency in (4.1.44) or by the group delay in (4.1.46), the analysis in the *separate* time or frequency domains is not sufficient to reveal the basic structure of the eigenfunctions, except for the simple cases of LTI or multiplicative

channels. It is thus understandable that a more appropriate tool for analyzing the channel eigenfunctions is provided by the *joint* time-frequency representations (in App.4.5.2 we recall a few fundamental properties of time-frequency distributions (TFD) and the interested reader can refer to [18] or [32], for more details). The basic TFD that we use here is the Wigner-Ville Distribution (WVD). However, the WVD alone would exhibit high inner and outer interference terms, due to its non-linear nature. To reduce the interference terms and still preserve a good resolution in time and frequency, we have used the so called Smoothed Pseudo-WVD with reassignment, or RSPWVD for short, introduced in [6]. The basic property that we wish to exploit here is that, if the signal has a constant amplitude within a certain time interval and it is zero outside, its RSPWVD is mainly concentrated along the instantaneous frequency curve $f = f(t)$ [6]. This property is then particularly useful to test the validity of ***P1*** $\div$ ***P3***. More specifically, to verify (4.1.44) and (4.1.45) we proceed numerically through the following steps: i) build the matrix $\boldsymbol{H}(n)$ of the equivalent DT channel, as in (4.1.28); ii) compute the singular value decomposition: $\boldsymbol{H}(n) = \boldsymbol{U}(n)\boldsymbol{\Lambda}(n)\boldsymbol{V}^H(n)$; iii) compute the RSPWVD of the singular vectors $\boldsymbol{v}_k(n)$ and $\boldsymbol{u}_k(n)$; iv) compare the contour lines of the RSPWVD of $\boldsymbol{v}_k(n)$ with the contour lines of $|H(t,f)|$ corresponding to the level $\lambda_k(n)$. If the Nyquist's sampling criterion is respected, the TFD of the singular vector is sufficiently representative of the corresponding continuous time eigenfunction.

Three-ray multipath channel

Let us consider a three-ray channel with amplitudes $[1, 0.9j, 0.8]$, delays $[2.25, 4.875, 6.125]T_s$ and Doppler shifts $[8, 8.325, 6.25]/MT_s$, with $M = 128$. The modulus of the time-varying transfer function is sketched in Fig. 4.1 (a). To check the validity of (4.1.44) and (4.1.45), we report in Fig. 4.1(b) the contour lines of the function $|H(t,f)|$, cut at levels λ_{17} and λ_{25}, and in Fig. 4.1(c) the RSPWVD of the singular vectors associated to the eigenvalues λ_{17} and λ_{25}. Comparing Figs. 4.1(b) and 4.1(c), we observe that indeed, as predicted by ***P1*** and ***P3***: i) the TFD's of the right singular vectors $\boldsymbol{v}_k$ associated to the singular values λ_k (right column) are mainly concentrated along the curves where $|H(t,f)|^2 = \lambda_k^2$; ii) the TFD's of the left singular vectors $\boldsymbol{u}_k$ are shifted replicas of the TFD's of the right singular vectors $\boldsymbol{v}_k$. Interestingly, even though in this example $\max_{p,q} |\tau_p \nu_q| \simeq 0.4$, the analytic model (4.1.44)-4.1.45) predicts the behavior of the channel eigenfunctions very well.

LTV bandpass channels in time and frequency

Let us now consider the time-varying channel described by the following input-output relationship

$$y(t) = \int_{-T}^{T} x(\tau) sinc(2\pi B(t-\tau)) d\tau. \tag{4.1.62}$$

This channel stops all the time components outside the window $[-T, T]$ and all the frequency components outside the window $[-B, B]$. It is known that the eigenfunctions of this channel are the prolate spheroidal wave functions [87], [55], which

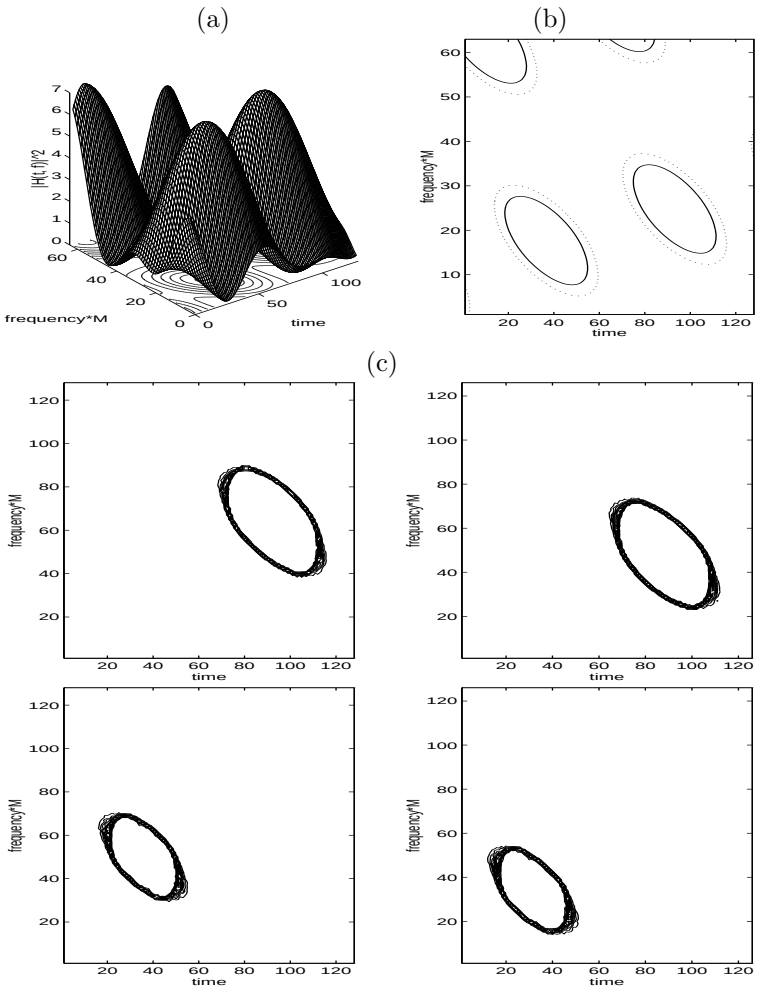

Figure 4.1. $|H(t,f)|^2$ of a 3-rays channel; (a) mesh plot; (b) contour plot of two different levels (singular values) of $\boldsymbol{H}$: λ_{17} (solid line) and λ_{25} (dashed line); (c) contour plot of the TFD of the left (left column) and right (right column) singular vectors associated to the singular values λ_{25} (upper row) and λ_{17} (lower row) of the matrix $\boldsymbol{H}$ describing the channel of Fig. 4.1.

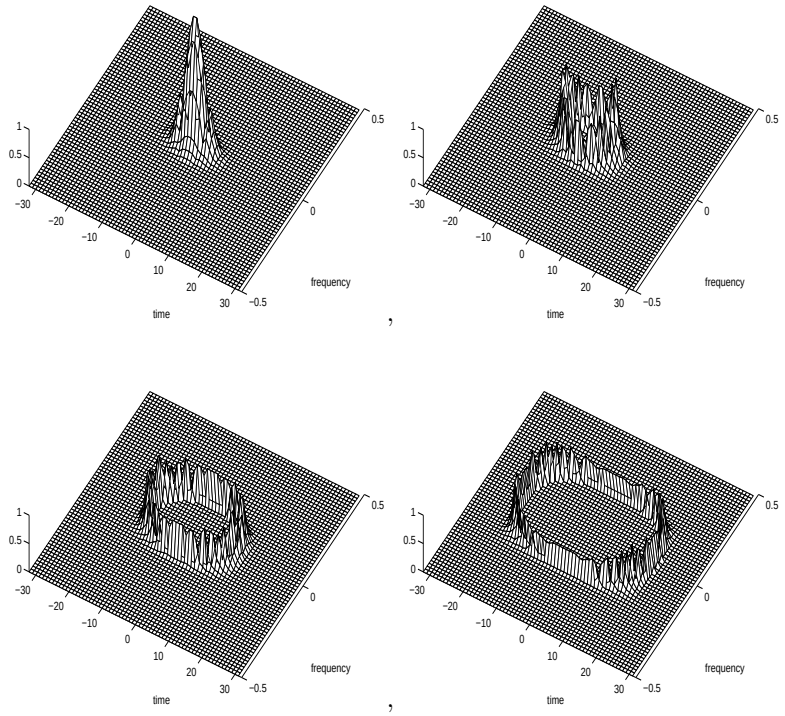

Figure 4.2. TFD's of the eigenfunctions of the channel (4.1.62) associated to the eigenvalues λ_1, λ_2, λ_4 and λ_8 (the reading order is from left to right, from top to bottom)

.

remarkably exhibit the same structure in time and frequency [69]. Hence, in this case the validity of ***P4*** can be proved analytically. Interestingly, the prolate spheroidal functions are also characterized by a behavior in the time-frequency domain similar to that observed in the previous example. As an example, let us assume that the channel allows only the transit of the components belonging to the time interval $[-24T_s, 24T_s]$ and to the frequency interval $[-0.25/T_s, 0.25/T_s]$. The RSPWVD's of some of the channel singular vectors (discrete-time prolate spheroidal functions), corresponding to the eigenvalues λ_1, λ_2, λ_4, and λ_8 are reported in Fig. 4.2. It is interesting to notice that the instantaneous frequencies (IF) are again closed curves in the time-frequency domain and the TFD's are highly concentrated along the IF curves, as in the previous example. In this example, the TFD's tend to fill the region between $-24T_s$ and $24T_s$ in time and between $-0.25/T_s$ and $0.25/T_s$ in frequency, i.e. the time and frequency passband of the system.

In these last two examples we observed that the instantaneous frequency describes a closed curve in the time-frequency plane giving rise to the rather peculiar *bubble* structure shown in Figs. 4.1 and 4.2. Interestingly, the bubble structure characterizes also the Hermite functions of high order. In fact, prolate spheroidal functions and Hermite polynomials are related by the following properties [31], [20]: i) the prolate spheroidal functions are the eigenfunctions of linear operators that maximize the signal energy concentration within rectangular regions of the time-frequency plane; ii) the Hermite functions are the eigenfunctions of linear operators that maximally concentrate the signal energy within a disk-shaped region of the time-frequency plane. Indeed, the bubble structure observed in the last examples, instead of being an exception, is rather the usual case because in general the contour lines of $|H(t, f)|$ are closed curves. Important exceptions arise though in very simple cases of LTV channels, such as: i) LTI channels, where the concentration curves are lines parallel to the time axis; ii) multiplicative channels, where the curves degenerate into lines parallel to the frequency axis; iii) the two-ray channel model where the curves become lines having an orientation depending on the channel delays and Doppler shifts. However, all these examples can be seen as ellipses degenerating into straight lines.

Before concluding this section, it is important to remark that, besides the expected greater complexity resulting from the derivation of the eigenfunctions of LTV channels, the eigenfunction of LTV channels do not share some of the fundamental properties of the eigenfunctions of LTI channels. Namely, i) except for the two limit cases of LTI and multiplicative channels, the behavior of the channel eigenfunctions is *channel-dependent*, as opposed to LTI systems whose eigenfunctions are always complex exponentials with constant amplitude and linear phase, irrespective of the channel; ii) furthermore, two cascaded LTV systems in general do not commute and the eigenfunctions of an LTV system composed of the cascade of two LTV subsystems are not simply related to the eigenfunctions of the two subsystems. However, the greater complexity of LTV systems is not necessarily a drawback, as long as their properties are properly taken into account and exploited advantageously, as we will see in the next sections.

4.2 Coding strategies for transmissions over LTV channels

In this section we show how to select the coding matrices $\boldsymbol{F}(n)$ and $\boldsymbol{G}(n)$ in (4.1.29) as a function of the channel status information (CSI), in order to optimize some performance parameter, such as output SNR or information rate. The general formulation provides a solution valid for any structure of the channel matrix $\boldsymbol{H}(n)$, so that the results have wider applicability than linear precoding for transmission over LTV channels. Specifically, we start in Section 4.2.1 assuming perfect CSI available at both transmit and receive sides. Then, exploiting the relationship between channel singular values and time-varying transfer function, established by ***P1*** $\div$ ***P3***, in Section 4.2.2 we show how different optimization strategies distribute the available transmit power (bits) as a function of time and frequency. In Section

4.2.3, we discuss some possible sub-optimal coding schemes resulting from a trade-off between performance and system complexity. Finally, in Section 4.2.4 we remove the assumption about perfect knowledge of the CSI and show how the uncertainties in our channel knowledge can be incorporated in the coding strategy design.

4.2.1 Perfect CSI available at both transmit and receive sides

In this section we assume perfect knowledge of the channel at both transmitter and receiver sides. Starting from the DT matrix model (4.1.29), we derive the optimal linear coding strategy according to different optimization criteria and constraints. We anticipate that all the solutions lead to a joint precoding/decoding strategy that decomposes the possibly time and frequency dispersive channel into multiple parallel uncorrelated non-dispersive subchannels. The distinction among the criteria is reflected only through the way power and bits are distributed across subchannels and blocks. Furthermore, in all cases, the columns of the coding matrix $\boldsymbol{F}(n)$ in (4.1.29), pertaining to the n-th transmitted block, are simply proportional to the right singular vectors of the channel matrix. This result generalizes the result obtained for LTI channels, analyzed in greater detail in Chapter 9 of Vol. I. However, besides the obvious difference due to the block-to-block variability of $\boldsymbol{H}(n)$, there is one more basic difference which is worth pointing out. The difference concerns the joint power/bits allocation across both subchannels *and* successive blocks, which is a unique degree of freedom offered by the channel variability.

In general, the transmitters operate with a prescribed average transmit power so that it is necessary to incorporate the average power constraint in the optimization strategy. With reference to (4.1.29), assuming that the information symbols are i.i.d. with zero mean and unit variance, the power constraint corresponds to

$$\frac{1}{N_b}\sum_{i=n}^{n+N_b-1} tr(\boldsymbol{F}(i)^H\boldsymbol{F}(i)) = \mathcal{P}_0, \tag{4.2.1}$$

where N_b is the number of blocks over which power allocation is optimized. Taking $N_b = 1$, we follow a *short-term* power allocation, whereas increasing N_b we have the possibility of distributing the power along successive blocks better by adopting a *long-term* allocation policy. The distinction between short and long term strategies has been analyzed in great detail in [15] for random channels exhibiting flat fading within each block, but varying from block to block. In principle, the best allocation of resources, i.e. power and bits, would require a high value of N_b (ideally infinite). In practice, N_b must result from a trade-off between system performance, maximum decoding delay and complexity.

Starting from the SVD of the channel matrix $\boldsymbol{H}(n) := \boldsymbol{U}(n)\boldsymbol{\Lambda}(n)\boldsymbol{V}^H(n)$ [c.f. (4.1.42)] and assuming that the symbols are uncorrelated and the noise is white and Gaussian, we show next that the matrices $\boldsymbol{F}_{opt}(n)$ and $\boldsymbol{G}_{opt}(n)$ resulting from all optimization criteria, can *always* be expressed as (see also Chapter 9 of Vol. I):

$$\boldsymbol{F}_{opt}(n) = \boldsymbol{V}(n)\boldsymbol{\Phi}(n) \quad , \quad \boldsymbol{G}_{opt}(n) = \boldsymbol{\Gamma}(n)\boldsymbol{\Lambda}^{-1}(n)\boldsymbol{U}^H(n) \tag{4.2.2}$$

where $\mathbf{\Phi}(n)$ and $\mathbf{\Gamma}(n)$ are diagonal matrices. Specifically, $\mathbf{\Phi}(n)$ determines the optimal power loading across subchannels in the nth block, whereas $\mathbf{\Gamma}(n)$ depends on the adopted equalization strategy. As an immediate consequence of (4.2.2), the channel linking the information symbol vector $\boldsymbol{s}(n)$ with the estimated vector $\hat{\boldsymbol{s}}(n)$ is characterized by the diagonal matrix $\boldsymbol{G}(n)\boldsymbol{H}(n)\boldsymbol{F}(n) = \mathbf{\Gamma}(n)\mathbf{\Phi}(n)$. Furthermore, the output noise vector $\boldsymbol{w}(n) = \boldsymbol{G}(n)\boldsymbol{v}(n)$ is composed of uncorrelated, and thus independent, samples because its covariance matrix $\boldsymbol{G}(n)\boldsymbol{G}^H(n)\sigma_v^2 = \mathbf{\Gamma}(n)\mathbf{\Lambda}^{-2}(n)\mathbf{\Gamma}^H(n)\sigma_v^2$ is also diagonal. Hence the overall I/O relationship is equivalent to a set of M independent flat fading subchannels:

$$\hat{s}_m(n) = \gamma_m(n)\phi_m(n)s_m(n) + \gamma_m(n)\beta_m(n) \tag{4.2.3}$$

where $\gamma_m(n)$ and $\phi_m(n)$ denote the diagonal entries of $\mathbf{\Gamma}(n)$ and $\mathbf{\Phi}(n)$, whereas $\beta_m(n)$ is additive noise with variance $\sigma_v^2\lambda_m^{-2}(n)$.

In the following, we specialize (4.2.2) to the following optimization criteria: i) minimum mean square error (MSE) between information symbols $\boldsymbol{s}(n)$ and estimated symbols $\hat{\boldsymbol{s}}(n)$ under a zero-forcing (ZF) or an average transmit power (AP) constraint and ii) maximum information rate, subject to the AP constraint.

Theorem 4.2.1 (MMSE-ZF). *The pair $\boldsymbol{F}_{opt}(n)$ and $\boldsymbol{G}_{opt}(n)$ that yields minimum MSE (MMSE) with a ZF receiver results from the solution of*

$$\min_{\boldsymbol{F}(n),\ \boldsymbol{G}(n)} E\{||\hat{\boldsymbol{s}}(n) - \boldsymbol{s}(n)||^2\} \quad \text{subject to} \quad \boldsymbol{G}(n)\boldsymbol{H}(n)\boldsymbol{F}(n) = \boldsymbol{I}, \tag{4.2.4}$$

and has the same structure as (4.2.2), with

$$\mathbf{\Phi}(n) = \varphi_n\mathbf{\Lambda}^{-1}(n) \qquad \mathbf{\Gamma}(n) = \mathbf{\Phi}(n)^{\dagger}. \tag{4.2.5}$$

Imposing the constraint (4.2.1), φ_n is such that

$$|\varphi_n|^2 = \frac{N_b\mathcal{P}_0}{\sum_{i=n}^{n+N_b-1}\sum_{k=1}^{M}\lambda_k^{-2}(i)}. \tag{4.2.6}$$

Theorem 4.2.2 (MMSE-AP). *The pair $\boldsymbol{F}_{opt}(n)$ and $\boldsymbol{G}_{opt}(n)$ that yields MMSE between the transmitted symbols and estimated symbols, subject to the average power constraint (4.2.1), is the solution of*

$$\min_{\boldsymbol{F}(n),\ \boldsymbol{G}(n)} E\{||\hat{\boldsymbol{s}}(n)-\boldsymbol{s}(n)||^2\} \quad \text{subject to} \quad \frac{1}{N_b}\sum_{i=n}^{n+N_b-1} tr(\boldsymbol{F}(i)^H\boldsymbol{F}(i)) = \mathcal{P}_0, \tag{4.2.7}$$

and is given by (4.2.2) with

$$|\phi_m(n)|^2 \;=\; \left[\frac{\bar{N}_b\mathcal{P}_0 + \sigma_v^2\sum_{i=n}^{n+\bar{N}_b-1}\sum_{k=1}^{M(n)}\lambda_k^{-2}(i)}{\sum_{i=n}^{n+\bar{N}_b-1}\sum_{k=1}^{M(n)}\lambda_k^{-1}(i)}\;\frac{1}{\lambda_m(n)} - \frac{\sigma_v^2}{\lambda_m^2(n)}\right]^+ \tag{4.2.8}$$

where $\bar{N}_b$ indicates the number of blocks where there is at least one value $|\phi_m(n)| > 0$, whereas $\bar{M}(n)$ indicates the number of active subchannels in the nth block, i.e. the number of channels with $\phi_m(n) > 0$, as a function of n; $\boldsymbol{\Gamma}(n) = \boldsymbol{\Phi}^H(n)(\boldsymbol{\Lambda}^{-2}(n) + \boldsymbol{\Phi}^H(n)\boldsymbol{\Phi}(n))^{-1}$.

Introducing the matrix $\boldsymbol{T}(n) := \boldsymbol{G}(n)\boldsymbol{H}(n)$ and the noise vector $\boldsymbol{w}(n) := \boldsymbol{G}(n)\boldsymbol{v}(n)$, the mutual information between the n-th transmitted data block $\boldsymbol{u}(n) = \boldsymbol{F}(n)\boldsymbol{s}(n)$ and decoded symbols block $\hat{\boldsymbol{s}}(n)$ blocks, conditioned to the channel $\boldsymbol{H}(n)$, is maximized for $\boldsymbol{u}(n)$ Gaussian and can be written as [2]:

$$I(\boldsymbol{u}, \hat{\boldsymbol{s}}; n) = \frac{1}{P} \log_2 |(\boldsymbol{R}_{uu}^{\dagger} + \boldsymbol{T}^H(n)\boldsymbol{R}_{ww}^{-1}\boldsymbol{T}(n))\boldsymbol{R}_{uu}| \ . \tag{4.2.9}$$

where $\boldsymbol{R}_{uu} = \boldsymbol{F}(n)\boldsymbol{R}_{ss}\boldsymbol{F}^H(n)$ and $\boldsymbol{R}_{ww} = \boldsymbol{G}(n)\boldsymbol{R}_{vv}\boldsymbol{G}^H(n)$ are the covariance matrices of $\boldsymbol{u}(n)$ and $\boldsymbol{w}(n)$ respectively. The maximum information rate (MIR) is achieved using the following (see also [64])

Theorem 4.2.3 (MIR-AP). *The pair $\boldsymbol{F}(n)$ and $\boldsymbol{G}(n)$ that maximizes the average information rate subject to the average power constraint solves*

$$\max_{\boldsymbol{F}(n),\ \boldsymbol{G}(n)} \sum_{i=n}^{n+N_b-1} I(\boldsymbol{u}, \hat{\boldsymbol{s}}; i) \qquad \text{subject to} \qquad \frac{1}{N_b} \sum_{i=n}^{n+N_b-1} tr(\boldsymbol{F}(i)^H \boldsymbol{F}(i)) = \mathcal{P}_0, \tag{4.2.10}$$

and is given by (4.2.2), where $\phi_m(n)$ is

$$|\phi_m(n)|^2 \ = \ \left[\frac{\bar{N}_b \mathcal{P}_0 + \sigma_v^2 \sum_{i=n}^{n+\bar{N}_b-1} \sum_{k=1}^{M(n)} \lambda_k^{-2}(i)}{M(n)\bar{N}_b} - \frac{\sigma_v^2}{\lambda_m^2(n)} \right]^+ . \tag{4.2.11}$$

$\boldsymbol{\Gamma}(n)$ *is diagonal and it can be either* $\boldsymbol{\Gamma}_{zf}(n) = \boldsymbol{\Phi}^{\dagger}(n)$, *to insures the zero-forcing condition, or* $\boldsymbol{\Gamma}_{mmse}(n) = \boldsymbol{\Phi}^H(n)\ (\boldsymbol{\Lambda}^{-2}(n) + \boldsymbol{\Phi}^H(n)\boldsymbol{\Phi}(n))^{-1}$, *to provide the MMSE solution. As in (4.2.8), $M(n)$is the number of active subchannels used in the transmission of the nth block.*

We derive now the expressions for the main system performance parameters, i.e. SNR, information rate, MSE and bit error rate (BER), resulting from the above optimization criteria. From (4.2.3), the output SNR on the mth subchannel of the nth block is always equal to

$$SNR_m(n) = |\phi_m(n)|^2 \frac{\lambda_m^2(n)}{\sigma_v^2}. \tag{4.2.12}$$

Then, depending on the adopted optimization criterion, the SNR assumes different distributions across subchannels and blocks. Specifically, we have

$$\{SNR_m(n)\}_{MMSE-ZF} = \frac{|\varphi_n|^2}{\sigma_v^2} = \frac{\bar{N}_b \mathcal{P}_0}{\sigma_v^2 \sum_{i=n}^{n+\bar{N}_b} \sum_{k=1}^{M} \lambda_k^{-2}(i)},$$

$$\{SNR_m(n)\}_{MMSE-AP} = \frac{\bar{N}_b \mathcal{P}_0/\sigma_v^2 + \sum_{i=n}^{n+\bar{N}_b} \sum_{k=1}^{\bar{M}(n)} \lambda_k^{-2}(i)}{\sum_{i=n}^{n+\bar{N}_b} \sum_{k=1}^{\bar{M}(n)} \lambda_k^{-1}(i)} \ \lambda_m(n) - 1,$$

$$\{SNR_m(n)\}_{MIR-AP} = \frac{\bar{N}_b \mathcal{P}_0/\sigma_v^2 + \sum_{i=n}^{n+\bar{N}_b} \sum_{k=1}^{\bar{M}(n)} \lambda_k^{-2}(i)}{\bar{M}(n)\bar{N}_b} \lambda_m^2(n) - 1. \quad (4.2.13)$$

The average mutual information pertaining to N_b blocks (4.2.9) is

$$I(\boldsymbol{u}; \hat{\boldsymbol{s}}) = \frac{1}{PN_b} \sum_{i=n}^{n+N_b-1} \sum_{m=1}^{M(i)} \log_2(1 + SNR_m(i)) \ . \quad (4.2.14)$$

Similarly, the output MSE can be derived as follows. Adopting the ZF solution $\boldsymbol{\Gamma}_{zf} = \boldsymbol{\Phi}^{\dagger}(n)$, the MSE relative to the nth block is

$$\mathcal{E}_{ZF}(\boldsymbol{\Phi}(n)) \ = \ \frac{1}{PN_b} \sum_{i=n}^{n+N_b-1} \sum_{m=1}^{M(i)} \frac{1}{SNR_m(i)}. \quad (4.2.15)$$

Otherwise, opting for the MMSE solution $\boldsymbol{\Gamma}_{mmse} = \boldsymbol{\Phi}^H(n)(\boldsymbol{\Lambda}^{-2}(n) + \boldsymbol{\Phi}^H(n)\boldsymbol{\Phi}(n))^{-1}$, the MSE is

$$\mathcal{E}_{MMSE}(\boldsymbol{\Phi}(n)) \ = \ \frac{1}{PN_b} \sum_{i=n}^{n+N_b-1} \sum_{m=1}^{M(i)} \frac{1}{1 + SNR_m(i)}. \quad (4.2.16)$$

Finally, even though a closed form expression for the BER is complicated, we can provide a simple upper bound. Because of the independence among subchannels guaranteed by the optimal coding, symbol decision on each subchannel is optimal and the overall BER is $\bar{P}_b = 1 - \prod_{i=n}^{n+\bar{N}_b-1} \prod_{m=1}^{M} [1 - P_m(i)]$, where $P_m(i)$ is the BER on the mth subchannel of the ith block. Assuming that $P_m(i) \ll 1, \forall m, i$, we have $\bar{P}_b \approx \sum_m \sum_i P_m(i)$. Furthermore, using a QAM constellation of size $Q_m(i)$ on the mth subchannel of the ith block and Gray encoding, the BER on each subchannel can be upper bounded as

$$P_m(i) < \frac{2}{\log_2 Q_m(i)} \ erfc\left(\sqrt{\frac{3SNR_m(i)}{2(Q_m(i) - 1)}}\right). \quad (4.2.17)$$

It is worth noticing how, differently from the MMSE/ZF strategy, the two strategies incorporating the average power constraint in the optimization prevent the transmission over the most attenuated subchannels, by setting the corresponding $\phi_m(n) = 0$ and distributing that power over the other subchannels. Further insight into the different power allocation strategies will be given in the ensuing section.

4.2.2 Comparisons and asymptotic bounds

In this section we evaluate the asymptotic expressions, valid as the block length M tends to infinity, of all the optimal power allocation strategies described in the

previous section. This formulation, even though approximate, is particularly important to shed light on the basic mechanisms controlling power and bit allocation as a function of time and frequency followed by each optimization strategy. It also suggests simpler sub-optimal implementations of the encoder, as shown in Section 4.2.3.

The asymptotic behaviors are derived by exploiting the relationship between the channel matrix singular values and the channel time-varying transfer function established by (4.1.45). As the blocklength M tends to infinity, the sequence λ_m of the multipath channel singular values converges towards a continuous function $\lambda(x)$ bounded between finite limits λ_{min} and λ_{max}. Since, according to (4.1.45), different values of λ map onto disjoint curves in the time-frequency plane, the integral of any function $\mathcal{F}(\lambda)$ over all values of λ can be computed, equivalently, as the integral of $\mathcal{F}(|H(t,f)|)$ over the time-frequency plane. In formulas, denoting by $\lambda_m^{(M)}$ the mth channel singular value in a block of length M (we drop the dependence on the block index n because, as M tends to infinity, there is only one block), asymptotically we have

$$\lim_{M\to\infty} \frac{1}{M} \sum_{m=1}^{M} \mathcal{F}(\lambda_m^{(M)}) \simeq \int_0^1 \mathcal{F}(\lambda(x))\ dx \simeq \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathcal{F}\left(|H(t,f)|\right)\ dt\ df, \tag{4.2.18}$$

where x has been normalized so that $x \in [0,1]$ and f has been normalized to the transmission bandwidth B.

Using the relationship (4.1.45), we can thus express all the basic system parameters, namely transmit power distribution, SNR, information rate and BER as a function of time and frequency. Specifically, substituting $|H(t,f)|$ to $\lambda_m(n)$ in (4.2.12), the SNR can be expressed as

$$SNR(t,f) = |\Phi(t,f)|^2 |H(t,f)|^2 / \sigma_v^2, \tag{4.2.19}$$

where $|\Phi(t,f)|^2$ depends on the adopted optimization criterion. In particular, from (4.2.5) the MMSE/ZF design leads to

$$|\Phi(t,f)|^2_{MMSE/ZF} = \frac{\mathcal{P}_0}{\sigma_v^2}\ \frac{|H(t,f)|^{-2}}{\int\int\ |H(\xi,\varphi)|^{-2}\ d\xi d\varphi} = \frac{\mathcal{K}_{ZF}}{|H(t,f)|^2}, \tag{4.2.20}$$

so that the final SNR turns out to be constant over the time-frequency plane:

$$SNR(t,f) \equiv \mathcal{K}_{ZF}. \tag{4.2.21}$$

Hence the power loading that yields the minimum MSE under the ZF constraint tends to equalize the SNR. However, this strategy may be highly inconvenient if the channel is strongly selective in time and/or in frequency, for most of the power would be wasted to compensate for the deepest channel nulls. Conversely, incorporating the average power constraint in the optimization method prevents such a potential waste of energy. In fact, according to both MMSE-AP or MIR-AP criteria, some

subchannels may be *inactive*, if $\mathcal{P}_0$ is insufficient to have all $|\phi_m(i)|^2 > 0$, for $m = 1, \ldots, M$ and $i = n, \ldots, n + \bar{N}_b - 1$. Equivalently, for $M \gg 1$ and using again (4.1.45), the inactive channels can be interpreted as regions of the time-frequency plane where no transmit power is allocated. These regions correspond to the (t, f) values where $\lambda^2 = |H(t, f)|^2$ falls below a certain value $\mathcal{T}(\mathcal{P}_0)$ which is the minimum value such that the function $|\Phi(t, f)|^2$ is strictly positive. Let us denote by $\mathcal{D}(\mathcal{P}_0)$ the region

$$\mathcal{D}(\mathcal{P}_0) := \{(t, f) \ : \ |H(t, f)|^2 \geq \mathcal{T}(\mathcal{P}_0)\}, \tag{4.2.22}$$

with $\mathcal{D}(\mathcal{P}_0)$ depending on $\mathcal{P}_0$ and on the optimization criterion. We are now able to write the power allocation laws for both MMSE/AP and MIR/AP criteria. Specifically, substituting $|H(t, f)|$ to λ in (4.2.8), the MMSE criterion loads power according to

$$\begin{aligned} |\Phi(t,f)|^2_{MMSE/AP}|_{(t,f)\in\mathcal{D}(\mathcal{P}_0)} &= \frac{\mathcal{P}_0/\sigma_v^2 + \int\int_{\mathcal{D}(\mathcal{P}_0)} |H(\xi,\varphi)|^{-2}\, d\xi d\varphi}{\int\int_{\mathcal{D}(\mathcal{P}_0)} |H(\xi,\varphi)|^{-1}\, d\xi d\varphi} \frac{1}{|H(t,f)|} - \frac{\sigma_v^2}{|H(t,f)|^2} \\ &:= \frac{\mathcal{K}_{MMSE}}{|H(t,f)|} - \frac{\sigma_v^2}{|H(t,f)|^2}. \end{aligned} \tag{4.2.23}$$

Similarly, from (4.2.11), the MIR-AP design leads to

$$\begin{aligned} |\Phi(t,f)|^2_{MIR/AP}\Big|_{(t,f)\in\mathcal{D}(\mathcal{P}_0)} &= \frac{\mathcal{P}_0/\sigma_v^2 + \int\int_{\mathcal{D}(\mathcal{P}_0)} |H(\xi,\varphi)|^{-2}\, d\xi d\varphi}{\int\int_{\mathcal{D}(\mathcal{P}_0)} d\xi d\varphi} - \frac{\sigma_v^2}{|H(t,f)|^2} \\ &:= \mathcal{K}_{MIR} - \frac{\sigma_v^2}{|H(t,f)|^2}. \end{aligned} \tag{4.2.24}$$

Interestingly, the MIR design implements the *water-pouring* strategy [33], extended to the time-frequency domain. A similar result was already suggested in [37], even though without an explicit analytic model of the channel eigenfunctions which, conversely, has been fundamental here to derive (4.2.24).

Observing the power allocation in (4.2.23), we may notice how the MMSE-AP criterion yields a power allocation which is in some way intermediate between the MMSE-ZF and the MIR-AP criteria, thus avoiding the two extreme behaviors (excessive power waste over the channel nulls for the MMSE-ZF method or most of the power allocated only over the best subchannels for the MIR-AP method).

Finally, recalling that the capacity of a deterministic channel is

$$C := \lim_{P\to\infty} \frac{1}{P} \max_{\boldsymbol{R}_{uu}} \boldsymbol{I}(\hat{\boldsymbol{s}}, \boldsymbol{u}), \tag{4.2.25}$$

where $\boldsymbol{R}_{uu}$ is the covariance matrix of the transmitted data vector $\boldsymbol{u}(n)$, the asymptotic optimal loading in (4.2.24) allows us to evaluate the channel capacity. From (4.2.14), (4.2.19) and (4.2.24), the capacity of a deterministic channel with additive

Gaussian noise is

$$C \approx \int\int_{\mathcal{D}(\mathcal{P}_0)} \log(1+SNR(t,f))dtdf \;=\; \int\int_{\mathcal{D}(\mathcal{P}_0)} \log\left(\mathcal{K}_{MIR}\frac{\sigma_v^2}{|H(t,f)|^2}\right)dtdf \tag{4.2.26}$$

This result confirms the validity of the expressions given in [63], [37].

Since the previous expressions are valid only asymptotically, as the block length tends to infinity, it is important to assess via numerical examples the block length beyond which the asymptotic limit is close to the result obtained with a finite block. In the next example we consider the information rate, but the same kind of conclusions can be drawn for the other cases as well. In particular, the asymptotic results are approached rather closely as soon as the block length exceeds the channel coherence time.

Example 4.2.1 (Capacity of deterministic LTV channels). *To verify the validity of (4.2.26), we compared the bound with the maximal information rate as a function of the block length. We considered the transmission with symbol rate $1/T_s$ over a three-ray channel, with amplitudes $h_0 = 1$, $h_1 = 0.9e^{j\pi/4}$, $h_2 = 0.9e^{-j\pi/4}$, Doppler frequencies $f_0 = 0$, $f_1 = 2.3/(64T_s)$, $f_2 = 1.4/(64T_s)$ and delays $\tau_0 = 0$, $\tau_1 = T_s$, $\tau_2 = 2T_s$. The SNR is 10 dB. The maximum mutual information given by (4.2.14), after power loading, is reported in Fig. 4.3 as a function of the block size. In particular, solid line refers to the LTV channel described above whereas the dashed line refers to the equivalent LTI channel, defined as the channel having the same parameters as the LTV channel, except for the Doppler frequencies which are equal to zero. The dotted line reports the capacity bound computed with (4.2.26). From Fig. 4.3 we can make the following remarks: i) the maximum average mutual information tends to a horizontal asymptote, which can be referred to as the channel capacity; ii) the maximum mutual information is approximately the same for the two equivalent LTI and LTV channels; iii) the bound (4.2.26) is rather close to the maximum information rate. Remark ii) indicates that the channel variability does not imply any loss on the maximum information rate, provided that proper coding is applied to the transmitted sequence or, in other words, that the transmitter is able to predict the channel evolution exactly on the basis of the CSI. In Fig. 4.4 we show the same quantities as in Fig. 4.3 obtained after averaging over 100 independent channel realizations (SNR = 4 dB). We considered a Rayleigh fading multipath channel with 4 rays (K=3) whose amplitudes h_k are generated as independent complex Gaussian random variables (normalized to have unit norm: $\sum_k |h_k|^2 = 1$). The sets of delays and Doppler frequencies are $(0, T_s, 2T_s, 3T_s)$ and $(0, -1.3/64,\ 3.5/64,\ 0.54/64)/T_s$, respectively. Also in this case, we observe a clear asymptotic behavior of the maximum information rate and the closeness of the capacity bound with the asymptote of the maximum information rate. We also observe that the information rate is very close to the capacity for block durations greater than the channel coherence time T_c (in this example $T_c \approx 26.6T_s$).*

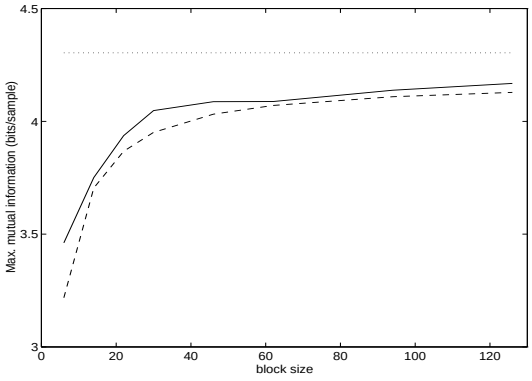

Figure 4.3. Capacity bound (dotted line) and maximum mutual information vs. block length for LTV (solid line) and LTI (dashed line) channels.

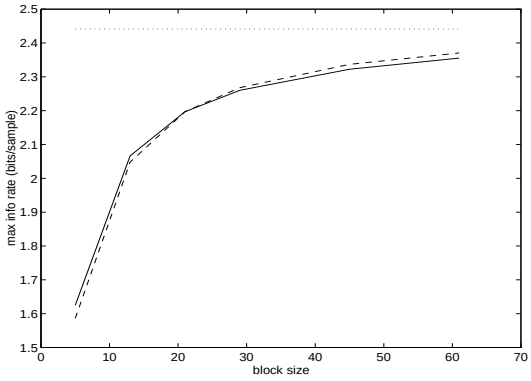

Figure 4.4. Capacity bound (dotted line) and maximum mutual information for LTV (solid line) and LTI (dashed line) channels, averaged over 100 channels.

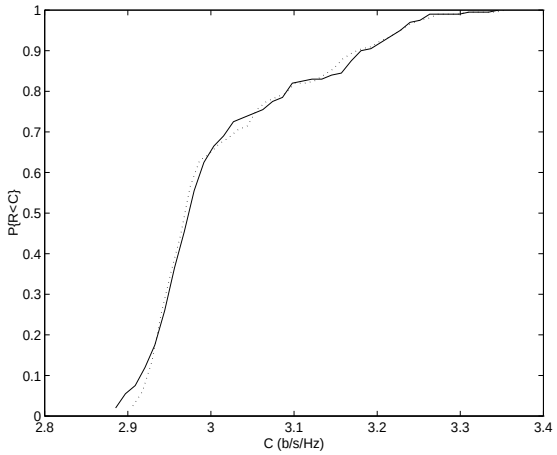

Figure 4.5. Sample cumulative distribution function of the maximal mutual information (solid line) and channel capacity (dotted line).

Information outage probability

As already mentioned in Section 4.1.2, a global characterization of the channel requires a stochastic channel modeling. As a consequence, interpreting every channel impulse response as a realization of a random process, the corresponding mutual information derived in (4.2.14) is a particular outcome of a random variable. Hence, a full channel characterization requires the computation of the pdf of the information rate. The so called information outage probability (IOP) has been proposed as a meaningful parameter for assessing the performance of link through a random LTV channel (see, e.g. [15]). The IOP is the probability that the mutual information exceeds a certain rate R. More specifically, in [15] the IOP was derived for flat fading block channels, where the channel is constant over a block, but varies randomly from block to block. The IOP was then minimized under different constraints on the transmitted power. The generalization of the results of [15] to a more general class of LTV channels is rather complicated. However, we will show next by means of a numerical example that the link between the singular values and the time-varying transfer function, established by (4.1.45) can be very helpful to find out an approximate behavior for the IOP's pdf.

Example 4.2.2. *: We assumed a multipath channel model, as in (4.1.13) with $Q = 3$, where the amplitudes h_q are complex uncorrelated Gaussian random variables with zero mean and unit variance, whereas the delay and Doppler values are fixed. In particular, the normalized delay vector is $\boldsymbol{\theta} = [0, 1, 2]$ and the normalized Doppler shift vector is $\boldsymbol{\nu} = [0, -0.735/64, 0.375/64]$. For each channel realization, we derived the DT channel model using (4.1.26), computed the SVD of the chan-*

nel matrix $\boldsymbol{H}(n)$ and the maximal information rate using (4.2.14). For the same realization, we computed the approximate channel capacity resulting from the CT model, using (4.2.26). We used long blocks (i.e. $P = 100$) so that the maximum information rate is very close to the channel capacity. In Fig. 4.5 we plot the sample cumulative distribution functions (cdf) of the information rate (solid line) and of the capacity (dotted line), estimated over 200 independent channel realizations. From Fig. (4.5) we can see that indeed the approximate relationship (4.1.45) between the channel singular values and the channel time-varying transfer function is very useful in deriving an approximate behavior of the information rate probability density function, without computing any SVD, but using only the function $|H(t, f)|$.

4.2.3 Adaptive OFDM

The implementation of the optimal coding may be problematic from the computational point of view because it requires the computation of an SVD for each block. Of course, due to the slow fluctuations of most practical channels, several simplifications are possible. However, it is clearly important to devise simple and fast sub-optimal coding schemes which do not require any SVD, to make the coding strategy appealing for practical applications, especially for services with a strict bound on the maximum decoding delay. For this reason, we suggest two possible suboptimal strategies.

a) Block time-varying OFDM (BTV-OFDM)

In the previous section we noticed that the optimal precoder's task is to distribute the available power (bits) as a function of both time and frequency. A simple sub-optimal way to implement this strategy is block time-varying OFDM (BTV-OFDM), where the channel is supposed to be (practically) time-invariant within a block, but it is time-varying from block to block. Intuitively speaking, BTV-OFDM assumes that the time-varying transfer function $H(t, f)$ is a piecewise constant function, constant over rectangles of dimension $PT_s \times 1/MT_s$ in the time-frequency plane. In such a case the channel matrix $\boldsymbol{H}(n)$ corresponding to each block is approximately Toeplitz and thus its SVD is

$$\boldsymbol{H}(n) \approx \boldsymbol{V}\boldsymbol{\Lambda}(n)\boldsymbol{U}^H \tag{4.2.27}$$

where the columns of $\boldsymbol{V}$ and $\boldsymbol{U}$ are orthonormal complex exponentials , irrespective of the channel status in block n. As a consequence, from (4.2.2) the columns of the optimal encoding matrix $\boldsymbol{F}$ are complex exponentials and thus the optimal coding strategy becomes equivalent to OFDM, except that the power loading varies from block to block. Clearly, the validity of this statement requires that the OFDM symbol duration $PT_s = (M+L)T_s$ is sufficiently smaller than the channel coherence time. Therefore the choice of P is critical because it is the only degree of freedom that we have to make this approximation more or less accurate. On one hand, if P is small, the approximation is good, but there is no advantage in using OFDM, because of the information rate loss due to the guard intervals; on the other hand,

larger values of P yield a better system efficiency, but make the assumption of $H(t,f)$ constant over PT_s seconds less accurate. Adaptive modulation schemes were proposed to achieve optimality using different criteria (see e.g. [16],[96],[4]), under a transmit power constraint and assuming the channel perfectly known at both receiver and transmitter sides. Here we consider the general case of channels dispersive in both time and frequency.

In OFDM, information symbols are parsed into consecutive M-long blocks $\boldsymbol{s}(n) := (s(nM), \ldots, s(nM+M-1))^T$, and mapped through a precoding matrix $\boldsymbol{F}$ onto the vector $\boldsymbol{u}(n) = \boldsymbol{F}\boldsymbol{s}(n)$ whose entries are modulated and transmitted through the channel. The columns of $\boldsymbol{F}$ are complex exponentials, $\{\boldsymbol{F}\}_{n,m} := \phi_m(n) e^{j\frac{2\pi}{M}m(n-L)}$ $m \in [0, M-1]$, with $M > L$, $n \in [0, P-1]$. Differently from standard OFDM, where $\phi_m(n)$ does not vary with n, in adaptive OFDM $\phi_m(n)$ is allowed to vary to load different powers across subchannels and blocks. Matrix $\boldsymbol{F}$ is tall ($(M+L)\times M$), for the rows of $\boldsymbol{F}$ are increased by the addition of a prefix or suffix of length L, greater than or equal to the channel order. Thus $\boldsymbol{u}(n) := (u(nP), \ldots, u(nP+P-1))^T$. Cyclic prefix or null guard intervals can be adopted. For the sake of clarity, here we assume the use of a cyclic prefix, but all the conclusions drawn in the following apply to the null guard interval case as well. At the receiver, after discarding the guard interval, IBI-free data blocks $\boldsymbol{y}(n) := (y(nP+L), \ldots, y(nP+P-1))^T$ are obtained. Since the basic assumption of BTV-OFDM is that the channel impulse response does not vary within one OFDM symbol, we set $h[nP+k, l] \approx h[nP, l]$, for $k = 0, \ldots, P-1$. Thus, the kth entry of $\boldsymbol{y}(n)$ is

$$\begin{aligned} y(nP+k) &\approx \sum_{l=0}^{L} h[nP,l] \sum_{m=0}^{M-1} \{\boldsymbol{F}\}_{k-l,m} \{\boldsymbol{s}(n)\}_m + v(nP+k) \\ &= \sum_{m=0}^{M-1} H(nP, f_m) e^{j\frac{2\pi m}{M}(k-L)} \phi_m(n) s_m(nM) + v(nP+k), \end{aligned} \tag{4.2.28}$$

for $k = L, \ldots, P-1$ where $f_m = m/M$, $m = 0, \ldots, M-1$, indicates a specific normalized frequency and $s_m(nM) := s(nM+m)$. Subsequently, an M point FFT is performed over $\boldsymbol{y}(n)$ and the μth FFT output sample is

$$Y(nP, f_\mu) := \sum_{k=L}^{P-1} e^{-j\frac{2\pi}{M}\mu k} y(nP+k) \approx H(nP, f_\mu)\phi_\mu(n) s_\mu(nM) + V(nP, f_\mu), \tag{4.2.29}$$

where $V(nP, f_\mu) := \sum_{k=L}^{P-1} \exp(-j\frac{2\pi}{M}\mu k) v(nP+k)$ are also AWGN samples for $\mu \in [0, M-1]$. The coefficients $|\phi_m(n)|$ control the power allocation across subchannels and along successive time blocks and can be evaluated according to different criteria. A simple yet effective way to compute $|\phi_m(n)|$, even though only approximately valid for underspread channels, consists in using the formulas derived in Section 4.2, substituting $\lambda_m(n)$ with $|H(nPT_s, m/MT_s)|$, as suggested by (4.1.45). In this way, no SVD's are required, and the only necessary information concerns

$H(t, f)$.

b) Chirped-OFDM

In BTV-OFDM, $|H(t, f)|$ is assumed to be constant within each block and varies from block to block. The validity of this assumption requires the selection of small values of the block size M, and this may induce considerable efficiency loss. To improve the efficiency it is thus necessary to increas M, but then the channel introduces inter-symbol interference (ISI). As we know from the previous sections, the only way to avoid ISI is to use the channel eigenfunctions instead of the complex sinusoids, but this requires SVD. However, exploiting (4.1.44) and (4.1.45), we can get an estimate of the channel eigenvectors, without taking any SVD. In fact, given $H(t, f)$, we can use (4.1.44) and (4.1.45) to derive the eigenfunctions and then, by sampling, the singular vectors of $\boldsymbol{H}(n)$. For small fluctuations of $\boldsymbol{H}(n)$, the instantaneous frequency $f_\lambda(t)$ is very slowly varying, so that the eigenfunctions are approximately sinusoids with a small frequency modulation (chirping), whose behavior follows the shape of $|H(t, f)|$. Using these chirp signals, matched to the channel time-varying transfer function, we can thus use longer blocks, to improve the efficiency, without computing any SVD and still avoiding ISI.

4.2.4 Coding with partial CSI

While in the previous sections the channel was assumed to be deterministic and perfectly known, here we assume, more realistically, that the channel is random and it is known only with a certain accuracy (estimation of the channel parameters is considered in the next section). Partial knowledge of the channel was included in [43], where the proposed adaptive modulation assumed that the channel was a Gaussian process (Rayleigh fading), so that the pdf of the channel at a certain instant, conditioned on its past, followed a Gaussian law with mean value equal to the MMSE channel linear prediction and variance equal to the prediction variance. While in [43] the channel correlation was assumed to be known, in [29] a similar approach was used for equalization but based on the channel correlation estimates, obtained by modeling the channel taps as AR processes and using spectral estimation techniques to retrieve the AR parameters. Both [43] and [29] assumed the channel to be stationary and *flat fading* within blocks of data. Very recently, the adaptive modulation idea was extended to time *and* frequency selective channels using OFDM (see e.g. [98]).

In this section, unlike [98] and similarly to [43], [29],[4], the uncertainty in channel knowledge is incorporated in the encoding design. We assume that the channel impulse response $h[n, l]$ is composed of correlated zero mean jointly Gaussian random variables and it is stationary both in n and l, i.e.

$$E\{h^*[p, k]h[p + n, k + l]\} = R_{hh}(n, l). \tag{4.2.30}$$

Assuming $h[nP+k,l] \approx h(nP,l)$, for $k=0,\ldots,P-1$, the channel vector is

$$\boldsymbol{h}(n) := (h[nP,0],\ldots,h[nP,L])^T \;\sim\; \mathcal{N}(\boldsymbol{0},\boldsymbol{R}_{hh}(0))$$

where $\mathcal{N}(\boldsymbol{\mu},\boldsymbol{C})$ indicates a multivariate complex Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{C}$, where $\boldsymbol{R}_{hh}(n)$ is defined as

$$\{\boldsymbol{R}_{hh}(n)\}_{k,l} := E\{h^*[qP,k]h[qP+nP,l]\} \tag{4.2.31}$$

Assuming that the channel impulse response $\boldsymbol{h}(n)$ is exactly known at the 0th block, the conditional pdf of $\boldsymbol{h}(n)$ given $\boldsymbol{h}(0)$ is $\mathcal{N}(\hat{\boldsymbol{h}}_{LP}(n),\boldsymbol{C}_{LP}(n))$ where:

$$\hat{\boldsymbol{h}}_{LP}(n) \;:=\; \boldsymbol{R}_{hh}(n)\boldsymbol{R}_{hh}^{-1}(0)\boldsymbol{h}(0), \tag{4.2.32}$$

$$\boldsymbol{C}_{LP}(n) \;:=\; \boldsymbol{R}_{hh}(0) - \boldsymbol{R}_{hh}(n)\boldsymbol{R}_{hh}^{-1}(0)\boldsymbol{R}_{hh}^{H}(n), \tag{4.2.33}$$

with $\hat{\boldsymbol{h}}_{LP}(n)$ indicating the linearly predicted channel impulse response and $\boldsymbol{C}_{LP}(n)$ its covariance. To incorporate the initial error on the knowledge of $\boldsymbol{h}(0)$, i.e. $\Delta\boldsymbol{h}(0) := \boldsymbol{h}(0) - \hat{\boldsymbol{h}}(0)$, we can also assume $\Delta\boldsymbol{h}(0) \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{C}_E(0))$ and independent of $\boldsymbol{h}(n)$.

From the pdf of $\boldsymbol{h}(n)$, because $H(nP,f_m) = \boldsymbol{e}_m^H\boldsymbol{h}(n)$ where $\boldsymbol{e}_m = (1,\exp(j2\pi f_m),\ldots,\exp(j2\pi f_m L))^T$, we can easily derive the pdf of $H(nP,f_m)$. In fact,

$$H(nP,f_m) \sim \mathcal{N}(\hat{H}_{LP}(nP,f_m),\sigma^2_{LP}(nP,f_m)) \tag{4.2.34}$$

where $\hat{H}_{LP}(nP,f_m) := \boldsymbol{e}_m^H\,\hat{\boldsymbol{h}}_{LP}(n)$ and $\sigma^2_{LP}(nP,f_m) := \boldsymbol{e}_m^H\boldsymbol{C}_{LP}(n)\boldsymbol{e}_m$.

We derive now a closed form solution for the optimal power loading maximizing the average SNR, subject to the average power constraint (4.2.1). In formulas, using the Lagrange multipliers, we seek the $\Phi_m(n)$ that maximize the functional

$$\mathcal{J}(\boldsymbol{\Phi}) := \frac{1}{N_b}\sum_{n=1}^{N_b}\sum_{m=0}^{M-1} SNR_m(n) - \mu\Big(\frac{1}{N_b}\sum_{n=1}^{N_b}\sum_{m=0}^{M-1}|\Phi_m(n)|^2 - \mathcal{P}_0\Big), \tag{4.2.35}$$

where $\boldsymbol{\Phi} := (\Phi_0(1),\ldots,\Phi_{M-1}(N_b))^T$ is the subcarrier amplitude matrix, $\mathcal{P}_0$ is the average power over N_b consecutive blocks and μ is the Lagrange multiplier; $SNR_m(n)$ is the signal to noise ratio in the mth subcarrier of the nth block, where the noise incorporates both thermal noise and prediction error:

$$SNR_m(n) = \frac{|\hat{H}_{LP}(nP,f_m)|^2|\Phi_m(n)|^2\sigma_s^2}{\sigma^2_{LP}(nP,f_m)|\Phi_m(n)|^2\sigma_s^2+\sigma_v^2} = \alpha_m(n)\left(1-\frac{1}{\rho_m(n)|\Phi_m(n)|^2+1}\right), \tag{4.2.36}$$

where

$$\alpha_m(n) := \frac{|\hat{H}_{LP}(nP,f_m)|^2}{\sigma^2_{LP}(nP,f_m)} \quad \textit{and} \quad \rho_m(n) := \frac{\sigma^2_{LP}(nP,f_m)\sigma_s^2}{\sigma_v^2}, \tag{4.2.37}$$

and σ_s^2 is the symbol variance. Setting to zeros the derivatives of $\mathcal{J}(\mathbf{\Phi})$ with respect to $\Phi_m(n)$ in (4.2.35), we get

$$\frac{\alpha_m(n)\rho_m(n)}{(\rho_m(n)|\Phi_m(n)|^2+1)^2}\Phi_m^*(n) - \lambda\Phi_m^*(n) = 0, \tag{4.2.38}$$

which implies either $\Phi_m(n) = 0$ or:

$$|\Phi_m(n)|^2 = \left(\sqrt{\frac{\alpha_m(n)\rho_m(n)}{\lambda}} - 1\right)\frac{1}{\rho_m(n)} \quad . \tag{4.2.39}$$

The values $|\Phi_m(n)|^2$ are ordered in decreasing sense. Then, starting from the smallest value of $|\Phi_m(n)|^2$, if $|\Phi_m(n)|^2 \leq 0$, the corresponding m-th subchannel in the nth block is eliminated and the power is redistributed across the remaining subchannels. This procedure is repeated until all remaining $|\Phi_m(n)|^2$ are strictly positive. Denoting by $M(n)$ the number of subchannels which are really used in the nth block, enforcing the constraint on the average power we get the Lagrange multiplier

$$\frac{1}{\sqrt{\lambda}} = \frac{N_b\mathcal{P}_0 + \sum_{n=1}^{N_b}\sum_{m=0}^{M(n)-1}\rho_m(n)^{-1}}{\sum_{n=1}^{N_b}\sum_{m=0}^{M(n)-1}\sqrt{\alpha_m(n)/\rho_m(n)}}. \tag{4.2.40}$$

Plugging (4.2.40) back in (4.2.39) provides the closed form expression for $|\Phi_m(n)|^2$. Substituting $|\Phi_m(n)|^2$ in (4.2.39) inside (4.2.36) we obtain the SNR in the mth subchannel of the nth block

$$SNR(nP, f_m) = \alpha_m(n)\left(1 - \sqrt{\frac{\lambda}{\alpha_m(n)\rho_m(n)}}\right). \tag{4.2.41}$$

To analyze the performance of the transmission strategy described above, we consider the case of uncorrelated scattering (US) where the channel covariance matrix is diagonal, i.e. $\boldsymbol{R}_{hh}(n) := \rho_t(nP)\ diag(\sigma_0^2, \ldots, \sigma_{L-1}^2)$ where $\rho_t(\tau)$ denotes the time correlation coefficient of each channel tap at lag τ (we assume for simplicity that the correlation coefficient function is the same for all taps) and σ_l^2 is the variance of the lth path. Using (4.2.32) and (4.2.33), we can derive the expected value and variance of $H(nP, f_m)$:

$$\hat{H}_{LP}(nP, f_m) \ = \ \boldsymbol{e}_m^H\hat{\boldsymbol{h}}_{LP}(n) = \rho_t(nP)H(0, f_m); \tag{4.2.42}$$

$$\sigma_{LP}^2(nP, f_m) \ = \ \boldsymbol{e}_m^H\boldsymbol{C}_{LP}(n)\boldsymbol{e}_m = (1-\rho_t^2(n))\sum_{l=0}^{L}\sigma_l^2. \tag{4.2.43}$$

From (4.2.43) we note that the US model leads to a $\sigma_{LP}^2(nP, f_m)$ which is constant across subchannels. Simple choices for $\rho_t(nP)$ are: i) exponential:

$$\rho_t(nP) = e^{-nP/Nc} \tag{4.2.44}$$

where $N_c := T_c/T_s$ is the coherence time normalized to the sampling rate $T_s \approx 1/B$, and B is the bandwidth; or ii), the isotropic scattering model in [45, p.31],

$$\rho_t(nP) = J_0(2\pi nP/N_c). \tag{4.2.45}$$

Example 4.2.3. *In this example we consider a possible application to a wideband wireless local area network (LAN). Specifically, we refer to the High Performance Radio LAN (HIPERLAN) system that provides short distance, high speed radio links using the 5.2 GHz or the 17.1 GHz frequency bands. We generated random impulse responses according to the US model, using the power delay profile named "Channel A" (Fig.4.6), chosen as a typical indoor multipath scenario for HIPERLAN/2 in [28], operating at 5.2 GHz, with* $B = 200\ MHz$ *(*$T_s = 5\ nsec$*). The channel order is* $L \approx 19$*, for the impulse response samples beyond the* $19th$ *are strongly attenuated. We used OFDM blocks of size* $M = 64$ *(*$P = M + L$*) and modeled* $R_t(nP)$ *as in (4.2.44), with coherence time of* $10\mu sec$*, which corresponds to* $N_c \approx 10^4$ *and approximately to* 100 *blocks. Fig. 4.7 shows a realization of the channel time-varying transfer function* $H(nP, f_m)$*.*

In Fig. 4.8 we report the $SNR(nP, f_m)$ *resulting from the optimal short term power loading in (4.2.39), obtained with* $\mathcal{P}_0 = M$*, for an average* $SNR = 10$ *dBs. Subchannels with* $SNR(n, f_m)$ *below the value necessary to achieve* $P_e(n, f_m) \leq 10^{-3}$ *for BPSK, are discarded. This strategy is appropriate for data transmissions that must guarantee BER values below a certain level. From Fig. 4.8 we notice that, as* n *and thus the prediction error increases, i) the power tends to be equally distributed across channels and ii) the number of channels able to guarantee the required BER decreases and thus the data rate decreases. Imposing a bound on the minimal data rate, from curves like Fig. 4.8 we are able to assess after how many blocks it is necessary to update the channel estimate for not incurring into severe data rate losses.*

4.3 Channel estimation and prediction

The compensation of the linear distortion introduced by the transit through LTV channels requires knowledge, and thus an estimate, of the channel impulse response. The estimation can be parametric or not and based on training sequences or blind. In this section we assume a parametric modeling, which is important to derive performance parameters useful to quantify both estimation accuracy and prediction capabilities. For transmissions over LTV channels, it is fundamental to address the following issues: i) describe the channel status information (CSI) parsimoniously to minimize the information to be fed back to the transmitter and the complexity of channel prediction; ii) define appropriate estimation procedures for LTV channels; iii) devise channel prediction strategies; iv) quantify the estimation accuracy; v) quantify the prediction error. Parametric modeling of the channel impulse response is particularly useful in this context because it addresses constructively all the issues raised above by providing the following advantages: i) even though the channel

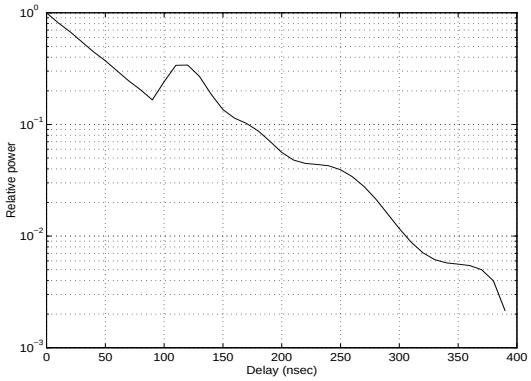

Figure 4.6. Channel power-delay profile.

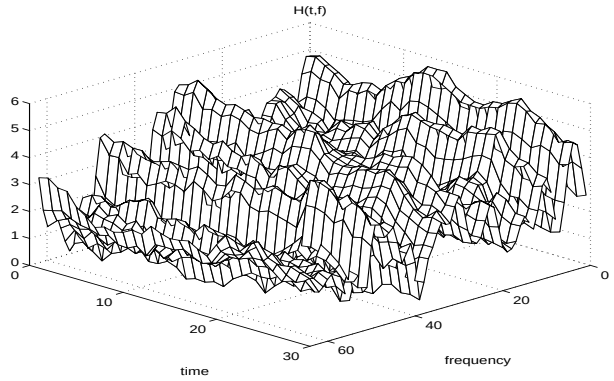

Figure 4.7. Time-varying transfer function.

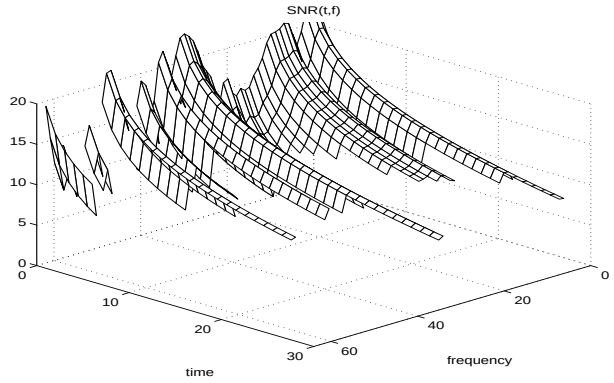

Figure 4.8. $SNR(n, f_m)$ resulting from optimal loading

impulse response is rapidly varying, its physical parameters, such as reflection coefficients, delays and Doppler shifts, vary much more slowly, so that transmitting the channel parameters rather than the impulse response samples is more efficient (see also [29]); ii) instead of estimating the coefficients of the DT impulse response, whose number is theoretically infinite, even for channels with finite delay, we can estimate the parameters of the CT channel, with a considerable reduction in number of unknowns; iii) the future channel evolution can be predicted by substituting the channel parameters' estimates in the channel model; iv) the bound on the channel parameters estimation accuracy can be evaluated using the Cramér-Rao bound (CRB), which is useful not only as a benchmark but also as a tool to quantify the impact of order underestimation or to compare blind versus training-based estimation techniques, independently of the adopted estimation algorithm; v) combining CRB and channel modeling, (approximate) closed form expressions can be derived for the channel prediction error, that determine the maximum time between successive updates of the channel estimates.

The channel parameter estimates are affected by two main error sources: i) observation noise and ii) mismatching between model and real channel. In this section we derive the theoretical limits for the accuracy of channel parameter estimates, considering both error sources. To provide a lower bound on the estimation accuracy and give a result independent of the estimation method, in Section 4.3.1 we derive the Cramér-Rao bound (CRB) for the channel parameters, valid for additive Gaussian noise channels. In Section 4.3.2, the CRB is then used to quantify the prediction error and finally, in Section 4.3.3 we suggest a simple method for estimating the parameters of a multipath LTV channel.

4.3.1 Cramér-Rao bound for LTV multipath channels

We derive the Cramér-Rao bound (CRB) for the accuracy of the LTV channel parameter estimates assuming the general multipath model (4.1.13) and additive white Gaussian noise. We consider, for the sake of brevity, only the case where the transmitted symbols are known (training-based estimation). However the analysis can be extended to blind estimators treating the transmitted symbols as nuisance parameters. We assume that IBI is completely removed thanks to the introduction of time guard intervals. The observed sequence is

$$y_k[n] = \sum_{l=0}^{M-1} h[nP+k, k-l]x_l[n] + v_k[n] \tag{4.3.1}$$

where $n = 0, \ldots, N-1$ is the block index, $k = 0, \ldots, P-1$ is the element index, $h[nP+k, l]$ is the DT channel impulse response given by (4.1.27), $x_l[n]$ is the lth coded symbol in the nth block and $\boldsymbol{v}(n) := (v_0[n], \ldots, v_{P-1}[n])^T \sim \mathcal{N}(\boldsymbol{0}, \sigma_v^2 \boldsymbol{I})$ is AGN. We assume the presence of a linear redundant precoder, so that the nth encoded block $\boldsymbol{x}(n) := (x_0(n), \ldots, x_{P-1}(n))^T$ is related to the corresponding information symbols block $\boldsymbol{s}(n) := (s_0(n), \ldots, s_{M-1}(n))^T$ through the $P \times M$ full column rank (complex) matrix $\boldsymbol{F}$ as $\boldsymbol{x}(n) = \boldsymbol{F}\boldsymbol{s}(n)$.

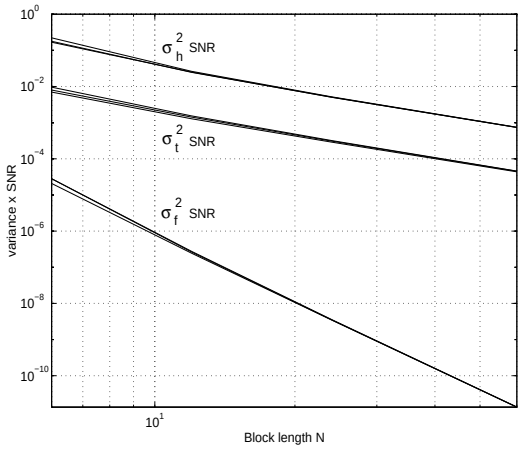

Figure 4.9. Cramér-Rao bounds in the estimate of channel amplitude (σ_h^2), delay (σ_t^2) and Doppler shifts (σ_f^2) vs. block length.

Under the AWGN assumption, the log-likelihood function relative to N consecutive blocks is then

$$f(\boldsymbol{\xi}) = c - \frac{1}{\sigma_v^2} \sum_{n=0}^{N-1} \sum_{k=0}^{P-1} |y_k[n] - \sum_{l=0}^{M-1} h[nP+k,l] x_l[n]|^2, \tag{4.3.2}$$

where the parameter vector $\boldsymbol{\xi} := (\boldsymbol{h}, \boldsymbol{\theta}, \boldsymbol{\nu})$ includes all channel parameters (for blind estimators, $\boldsymbol{\xi}$ contains also the transmitted symbols $x_k[n]$); c is a constant independent of the channel parameters. The derivation of the Fisher's information matrix elements is reported in the appendix. Here we report a few basic results to show the dependence of the estimation accuracy on SNR, block length, and number of blocks used for the estimation.

Estimation variance vs. N

In Fig. 4.9 we report the variance obtained in the estimate of the channel amplitudes h_q, normalized delays $\theta_q := \tau_q / T_s$ and Doppler shifts $\nu_q := f_q T_s$, as a function of the block length $N = M + L$, for a second order (i.e. $L = 2$) channel with parameters: $\boldsymbol{h} = [1, 1, 1]$, $\boldsymbol{\theta} = [0.83,\ 1.45,\ 2.1]$, and $\boldsymbol{\nu} = [0, 1.5/M, -1.5/M]$.

To get rid of the dependence on the specific training sequence used for the estimate, we averaged the results over several independent sequences, generated as vectors of i.i.d. QPSK symbols. The training is based on transmission of P-long blocks and the number of blocks used is $N = P$, for a total number of training symbols of $P^2 = N^2$. All the variances are inversely proportional to the SNR (the curves in Fig. 4.9 are multiplied by the SNR, to remove the dependence from the common factor $1/SNR$). The plots in Fig. 4.9 are shown in a log-log scale to

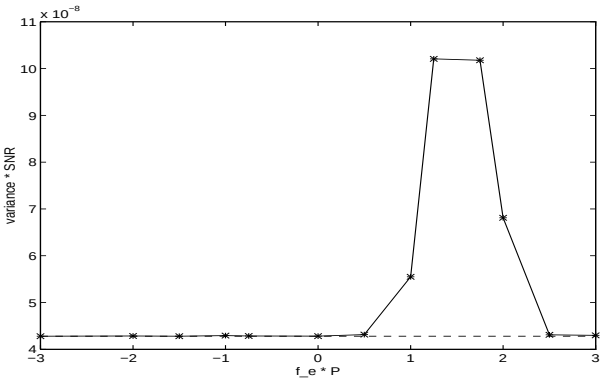

Figure 4.10. Variance in the estimate of f_2 as a function of the frequency f_e of the extra path.

evidence the power law behavior of the estimation variances. More specifically, the variance of both amplitude and delay estimates decreases as $1/N^2$, whereas the variance in the estimate of the Doppler frequency decreases as $1/N^6$. Considering that $N = P$, this means that the variances of the amplitude and delay estimates are inversely proportional to the number of samples (equal to N^2), whereas the variance in frequency estimate is inversely proportional to the cube of the number of samples.

Estimation variance in the presence of channel order underestimation

Most channel estimators are only able to track the principal components. It is then important to assess how the accuracy in the estimation of the dominant paths parameters is affected by the presence of extra undetected paths. The analysis of this situation can be performed again using the CRB, treating the extra paths parameters as nuisance parameters. As an example, we considered a multipath channel composed of four paths having the following parameters: $\boldsymbol{h} = [1, 1, 1, 0.25]$, $\boldsymbol{\theta} = [0.83, 1.45, 2.1, 1.45]$, and $\boldsymbol{\nu} = [0, 1.5/P, -1.5/P, f_e]$. We assume that the channel order estimate is $L = 3$, instead of 4, so that there is an extra path (the fourth one, which is also the weakest one). We fixed the delay of the extra path equal to the delay of the second path, so that the two paths interfere. Then we changed the frequency f_e of the extra path and analyzed how the variance in the estimate of the second path is affected by f_e. We considered, for simplicity only the variance of the frequency estimate because, as we will see in the next section, the frequency estimate is the most critical one, as far as the channel prediction is concerned. In Fig. 4.10 we show the variance of f_2 as a function of f_e. Of course, since $\tau_2 = \tau_4$, if f_e coincides with f_2, the Fisher Information matrix is not full rank and we cannot compute $\sigma^2_{f_2}$. However, it is interesting to observe the behavior of $\sigma^2_{f_2}$ as a function of f_e. With reference to Fig. 4.10 we observe that: i) the variance of f_2 increases

as f_e approaches f_2 (i.e. $1.5/P$), and ii) the estimate of f_2 is nearly unaffected by the fourth path as long as f_e differs from f_2 by more than the frequency resolution (i.e. $1/P$). On the basis of the result shown in Fig. 4.10, we may infer that clustered multipath components, i.e. paths whose parameters differ from each other for less than the time-frequency resolution, may cause considerable performance losses in case they are not detected, for they can affect the other parameters estimates considerably. Fortunately, from the channel prediction point of view, the situation is not so critical because these clustered paths are characterized by a very similar behavior and then the channel prediction error resulting from the undetected paths is less relevant than errors due to components which are well separated in time and/or frequency.

4.3.2 Channel prediction

The estimation of the channel parameters can be used to predict the channel time evolution inserting the channel parameter estimates in (4.1.26). Of course, since the estimates are inevitably affected by errors, it is reasonable to expect a channel prediction error that increases with the increase of the duration of the prediction interval, so that periodic updating is necessary. The aim of this section is to quantify the prediction error as a function of the block index, starting from the moment where the estimate is performed. More specifically, denoting by $\hat{h}_q$, $\hat{\theta}_q$ and $\hat{\nu}_q$ the channel estimates, with $q = 0, \ldots, Q-1$, the channel impulse response can be predicted inserting the parameter estimates in (4.1.27)

$$\hat{h}[nP+k,l] = \sum_{q=0}^{Q-1} \hat{h}_q e^{j\pi\hat{\nu}_q(2n-k+\hat{\theta}_q)} g[\pi(k-l-\hat{\theta}_q)] \tag{4.3.3}$$

where we set, for simplicity of notation $g(x) := sinc(x)$. Taking a first order approximation of the error, we get an approximate error expression valid asymptotically, for high SNR:

$$\begin{aligned}\epsilon[nP+k,l] \;:=\; \hat{h}[nP+k,l] - h[nP+k,l] \approx \sum_{q=0}^{Q-1} e^{j\pi\nu_q(nP+k+l+\theta_q)} \\ \cdot\;\; \{\epsilon_{h_q} g(\pi(k-l-\theta_q)) + \pi h_q \epsilon_{\theta_q}[j\nu_q g(\pi(k-l-\theta_q)) - \dot{g}(\pi(k-l-\theta_q))] \\ +\;\; j\pi h_q \epsilon_{\nu_q}(nP+k+l+\theta_q) g(\pi(k-l-\theta_q))\},\end{aligned} \tag{4.3.4}$$

where ϵ_{h_q}, ϵ_{θ_q} and ϵ_{ν_q} represent the errors in the estimate of the amplitude, delay and Doppler frequency of the qth channel path, respectively. The MSE can thus be computed as a function of the block index n, as

$$\begin{aligned}MSE(n) \;:= E\{|\epsilon[nP+k,l]|^2\} = \sum_{p=0}^{Q-1}\sum_{q=0}^{Q-1} e^{j\pi(\nu_p(nP+k+l+\theta_p) - \nu_q(nP+k+l+\theta_q))} \\ \cdot\;\; \Big\{ cov(\hat{h}_p, \hat{h}_q)\; g(\pi(k-l-\theta_p)) g(\pi(k-l-\theta_q))\end{aligned}$$

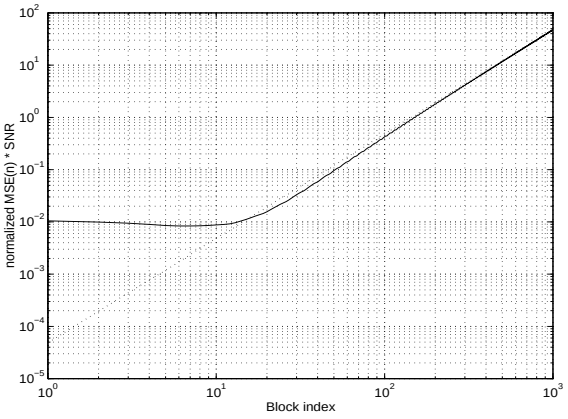

Figure 4.11. Normalized prediction error vs. block index n.

$$
\begin{aligned}
&+\ 2\pi\mathcal{R}e\left[cov(\hat{h}_p,\hat{\theta}_q)\ h_q^*\ g(\pi(k-l-\theta_p))[-j\nu_q g(\pi(k-l-\theta_q))-\dot{g}(\pi(k-l-\theta_q))]\right]\\
&-\ 2\pi\ \mathcal{R}e\left[cov(\hat{h}_p,\hat{\nu}_q)\ jh_q^*\ g(\pi(k-l-\theta_p))\ g(\pi(k-l-\theta_q))\ (nP+k+l+\theta_q)\right]\\
&+\ 2\pi^2\mathcal{R}e\ \Big[cov(\hat{\theta}_p,\hat{\theta}_q)\ h_p\ h_q^*\ (\nu_p\nu_q\ g(\pi(k-l-\theta_p))\ g(\pi(k-l-\theta_q))\\
&+\ \dot{g}(\pi(k-l-\theta_q)\ \dot{g}(\pi(k-l-\theta_q))-\nu_p\ g(\pi(k-l-\theta_p))\ \dot{g}(\pi(k-l-\theta_q))\\
&+\ j\nu_q\ \dot{g}(\pi(k-l-\theta_p))\ g(\pi(k-l-\theta_q)))\ \Big]\\
&+\ 2\pi^2\ \mathcal{R}e\left[cov(\hat{\theta}_p,\hat{\nu}_q)h_p h_q^*[\nu_p(nP+k+l+\theta_q)\ g(\pi(k-l-\theta_p))\ g(\pi(k-l-\theta_q))\right.\\
&+\ \left.j(nP+k+l+\theta_q)\ \dot{g}(\pi(k-l-\theta_p)\ g(\pi(k-l-\theta_q)))\ \right]\\
&+\ \pi^2\mathcal{R}e\left[cov(\hat{\nu}_p,\hat{\nu}_q)\ h_p h_q^*\ (nP+k+l+\theta_q)(nP+k+l+\theta_q)\right.\\
&\cdot\ \left.g(\pi(k-l-\theta_p))\ g(\pi(k-l-\theta_q))\ \right]\Big\}. \qquad (4.3.5)
\end{aligned}
$$

Thus, substituting the elements of the error covariance matrix, as given by the CRB for example, in (4.3.5), we can get MSE(n).

Example 4.3.1. *In Fig. 4.11 we report the normalized mean square error* $MSE(n)\cdot SNR/||\boldsymbol{h}(n)||^2$ *(solid line) as a function of the block index. The channel parameters are:* $\boldsymbol{h}=[1,0.9,0.8]$, $\boldsymbol{\theta}=[0,1.25,2.375]$; $\boldsymbol{\nu}=[0,1.5/M,-1.5/M]$. *In the same figure we report (dotted line) the contribution due to the last term in (4.3.5), to show that, as the block index increases, i) the main contribution to the MSE is given by the error in the frequency estimate; ii) at high* n, $MSE(n)$ *goes approximately as* n^2.

For high values of n, the dominant contribution of (4.3.5) is the last term. Therefore, at least asymptotically, i.e. for high values of n, we can approximate (4.3.5)

with its last term. Furthermore, to provide a result not constrained to a specific channel realization, we average (4.3.5) with respect to the channel amplitudes h_q. For example, assuming that the amplitudes h_q are uncorrelated random variables, with zero mean and the same variance σ_h^2, we can approximate the average value of (4.3.5), for high n, as follows

$$MSE(n) \simeq \sigma_h^2(nP) \sum_{p=0}^{Q-1} g^2(\pi(k-l-\theta_p))\sigma_{\hat{\nu}_p}^2. \tag{4.3.6}$$

This expression can be used to determine the training period. In fact, for a given maximum tolerable value of $MSE(n)$, (4.3.6) allows us to determine the maximum value of n beyond which the prediction error becomes intolerably high: That value of n determines the period of the estimation update.

4.3.3 Channel parameter estimation

Before concluding this section, we propose a simple method for block synchronization and for the estimation of the channel parameters based on the periodic transmission of training sequences. We do not consider blind estimation here, but the interested reader may refer to [34] or [90], for example. In contrast to most common communication systems, where the training sequences are pseudo-noise and the channel estimators seek the samples of the channel impulse response, here we use as training sequences linear frequency modulation signals (chirp signals) and estimate the channel parameters, i.e. delay and Doppler shifts, directly. Recalling (4.1.26), this implies a much more parsimonious approach in terms of number of unknowns. In fact, considering the Q triplets (h_q, τ_q, f_q), we have $3Q$ unknowns, in comparison with the DT impulse response which contains, theoretically, an infinite number of unknowns.

Our estimation method is as follows: At the beginning of each frame we send two consecutive blocks of length T containing the baseband chirp signals $s_1(t) = \exp(j\pi\mu t^2)$ and $s_2(t) = \exp(-j\pi\mu t^2)$ respectively. The two blocks are separated in time by a null guard interval whose duration T_g is greater than the channel memory. At the receiver, we multiply the baseband received signal by $s_1^*(t)$ in the first block and by $s_2^*(t)$ in the successive block. If the receiver is block synchronous with the transmitter, in the two successive intervals we receive the two signals

$$\begin{aligned} x_1(t) &= \sum_{q=0}^{Q-1} h_q e^{j\pi\mu(t-\tau_q)^2} e^{j2\pi f_q t} + v_1(t), t \in [0, T) \\ x_2(t) &= \sum_{q=0}^{Q-1} h_q e^{-j\pi\mu(t-\tau_q)^2} e^{j2\pi f_q t} + v_2(t), t \in [T+T_g, 2T+T_g) \end{aligned} \tag{4.3.7}$$

where $v_k(t)$ is additive noise in the interval k. After beating the received signals

with the conjugate of the transmitted chirp, we get

$$y_1(t) := x_1(t)e^{-j\pi\mu t^2} = \sum_{q=0}^{Q-1} h_q^{(1)} e^{j2\pi(f_q-\mu\tau_q)t} + w_1(t), t \in [0,T) \tag{4.3.8}$$

$$y_2(t) := x_2(t)e^{j\pi\mu t^2} = \sum_{q=0}^{Q-1} h_q^{(2)} e^{j2\pi(f_q+\mu\tau_q)t} + w_2(t), t \in [T+T_g, 2T+T_g)$$

where $h_q^{(1)} = h_q e^{j\pi\mu\tau_q^2}$, $h_q^{(2)} = h_q e^{-j\pi\mu\tau_q^2}$, $w_1(t) = v_1(t)e^{-j\pi\mu t^2}$ and $w_2(t) = v_2(t)e^{j\pi\mu t^2}$. Therefore, in the first block we have a signal composed of sinusoids of frequency $f_q^{(1)} = f_q - \mu\tau_q$, whereas in the second interval we get sinusoids of frequency $f_q^{(2)} = f_q + \mu\tau_q$. Thus equations (4.3.8) allow us to cast the channel estimation problem in terms of a harmonics retrieval problem. More specifically, the overall channel estimation method proceeds through the following steps:

1. Estimate complex amplitudes $\hat{h}_q^{(k)}$ and frequencies $\hat{f}_q^{(k)}$, for $q = 0, \ldots, Q-1$, of the sinusoids contained in $y_k(t)$, with $k = 1, 2$ using any harmonic retrieval algorithm (see, e.g. [88]);

2. Pair the sets of estimates obtained in the two blocks using the following procedure. Denoting by $(k_0^{(i)}, \ldots, k_{Q-1}^{(i)})$ the ith permutation of the index set $(0, 1, \ldots, Q-1)$, for each index permutation index i, with $i = 1, \ldots Q!$, estimate the channel Doppler frequencies and delays as

$$\begin{aligned} \hat{f}_{q,i} &= (\hat{f}_q^{(1)} + \hat{f}_{k_q^{(i)}}^{(2)})/2 \\ \hat{\tau}_{q,i} &= (\hat{f}_{k_q^{(i)}}^{(2)} - \hat{f}_q^{(1)})/2\mu, \quad i =, \ldots, Q!; \end{aligned} \tag{4.3.9}$$

3. Choose the index set that yields the MMSE, i.e.

$$\begin{aligned} i_{opt} = argmin\{ &\int_0^T |x_1(t) - \sum_{q=0}^{Q-1} h_q e^{j\pi\mu(t-\hat{\tau}_{q,i})^2} e^{j2\pi\hat{f}_{q,i}t}|^2 dt \\ &+ \int_{T+T_g}^{2T+T_g} |x_2(t) - \sum_{q=0}^{Q-1} h_q e^{-j\pi\mu(t-\hat{\tau}_{q,i})^2} e^{j2\pi\hat{f}_{q,i}t}|^2 dt\} \end{aligned} \tag{4.3.10}$$

and thus gives the estimates $\hat{f}_{q,i_{opt}}, \hat{\tau}_{q,i_{opt}}$, $q = 0, \ldots, Q-1$.

An example of application of this estimation method is reported in Fig. 4.12 where we show the variance in the estimation of amplitude, delay and Doppler shift as a function of the number of samples N used for the estimate. The channel is composed of three paths and the SNR is 15 dB. Each curve is the average value of the variances obtained in the estimates of homologous parameters. From Fig. 4.12 it is interesting to observe that the average variance in the estimate of the Doppler frequency decreases as $1/N^3$, whereas the average variances of amplitude and delay decrease as $1/N$, as with the CRB. Clearly the above method has at least a 3 dB loss due to the linear combination in (4.3.9), but it is also very simple, at least if Q is not too big to make the pairing operation troublesome.

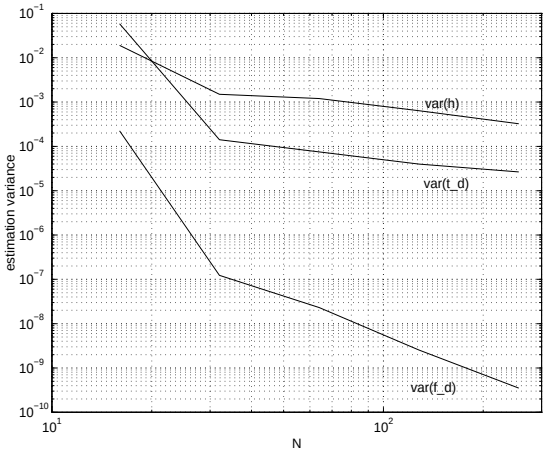

Figure 4.12. Average variances vs. number of samples N, in the estimate of amplitude, delay and Doppler shift of a multipath channel, using the chirp-based estimation algorithm (SNR=15 dB).

4.4 Conclusion

In this chapter we have addressed a few issues related to LTV systems and to the transmissions through LTV channels. We have proposed an approximate analytic model for the eigenfunctions of underspread LTV channels which, even if valid only in approximate sense, has been useful to envisage suboptimal transmission strategies and to predict the system performance in asymptotic sense. The joint time-frequency analysis has played an important role in modeling the eigenfunctions of linear time-varying channels. Interestingly, the theory developed in this chapter, valid for channels whose spread function is concentrated along the delay or the Doppler frequency axis, can be generalized to spread functions concentrated along curves of the (ν, τ) plane. In that case, the relationship between left and right singular functions becomes more complicated than in ***P3***, because the phase Ψ_0 in ***P3*** becomes also a function of time, i.e. $\Psi_0 = \Psi_0(t)$. Clearly the search for accurate models of the singular functions of LTV systems is one of the most challenging research areas, whose impact is not limited to transmissions over LTV channels.

The other main point addressed in this chapter is parametric modeling of the channel impulse response. The multipath channel parametric model has been used to derive simple methods for estimating and predicting the channel impulse response. But clearly, as with any parametric approach, the method lacks of robustness against mismatching between the real channel and the assumed model.

Probably the most interesting result in the study of LTV channels is that one realizes that there are so many research areas which are still in their embryonal state and need to be fully developed if one wants to exploit the potentialities intrinsic in

the transmission through LTV channels. We try here to suggest a few possible directions: i) simple yet reliable channel modeling, possibly fitting real data records; ii) robust channel estimation and prediction methods; iii) simple suboptimal adaptive methods for optimizing the transmission strategy, resulting from a good compromise between performance and complexity and incorporating the uncertainty in our channel knowledge; iv) derivation of system performance parameters such as outage probability or information outage probability for time and frequency dispersive stochastic channels; v) channel equalization or, more generally, reliable symbol detection algorithms incorporating channel fluctuations.

4.5 Appendices

4.5.1 Eigenfunction model

In this appendix we prove properties ***P1*** $\div$ ***P4*** assuming the multipath model (4.1.13). Nonetheless, since we allow Q (4.1.13) to go to infinity, the proof has wider validity. In fact, for any finite interval, in the time frequency plane (t, f), of size $\Delta t \times \Delta f$, the transfer function $H(t, f)$ can be expanded in Fourier series (the expansion requires only that $H(t, f)$ satisfies Dirichlet's condition in the finite interval) as

$$H(t,f) = \sum_{k=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} H_{k,m} e^{j2\pi \frac{k}{\Delta t} t} e^{-j2\pi \frac{m}{\Delta f} f}. \tag{4.5.1}$$

The corresponding impulse response is thus

$$h(t,\tau) = \sum_{k=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} H_{k,m} e^{j2\pi \frac{k}{\Delta t} t} \delta(\tau - \frac{m}{\Delta f}). \tag{4.5.2}$$

Since in the following proof we do not make any assumption about the summation indices in (4.1.13), we can consider (4.1.13) fairly general to include also (4.5.2) as a particular case. We adopt the multipath model here only to provide a physical insight in the approximations made during the proof. Furthermore, we assume initially that the spread function is centered around the origin. At the end of the proof we remove this assumption.

We start proving that the eigenfunction model given by (4.1.45) and (4.1.44) is valid, in approximate sense, for underspread channels. From the underspread assumption $\max_{p,q} |\tau_p \nu_q| \ll 1$, it is evident that we may have i) $\max_q |\tau_q| \ll 1$, but not necessarily $\max_q |\nu_q| \ll 1$ (frequency dispersive channel) ii) $\max_q |\nu_q| \ll 1$, but not necessarily $\max_q |\tau_q| \ll 1$ (time dispersive channel); iii) $\max_q |\nu_q| \ll 1$ and $\max_q |\tau_q| \ll 1$. The last case is the least interesting from the point of view of its eigenfunctions, because it represents an almost flat fading channel. This is a limiting case where all the eigenvalues tend to coincide with one single value given by the flat fading coefficient. Thus we concentrate our attention on the first two cases. In particular, from the mathematical point of view we can study only the first case, because, the second case can be analyzed by duality translating the previous approach in the frequency domain.

The composed kernel $\tilde{h}(\tau,\theta)$, given by (4.1.40), corresponding to the multipath channel (4.1.13), is

$$\tilde{h}(\tau,\theta)=\sum_p\sum_q h_p^* h_q e^{-j2\pi(\nu_p-\nu_q)(\tau+\tau_p)}\delta(\tau-\theta-\tau_q+\tau_p). \tag{4.5.3}$$

Substituting (4.5.3) in (4.1.39), the right singular function $v_\lambda(t)$ must be solution of the following equation

$$\lambda^2 v_\lambda(t)=\sum_p\sum_q h_p^* h_q e^{-j2\pi(\nu_p-\nu_q)(t+\tau_p)}v_\lambda(t-\tau_q+\tau_p). \tag{4.5.4}$$

We show next that an approximate solution of this equation, valid for time dispersive channels but with small dispersion in frequency, assumes the following form

$$v_\lambda(t)=\frac{1}{\sqrt{T_\lambda}}rect_{T_\lambda}(t-\xi_\lambda)e^{j\phi_\lambda(t)}, \tag{4.5.5}$$

where $\phi_\lambda(t)$ satisfies (4.1.45). For channels with $max_q|\nu_q|\ll 1$, T_λ is much greater than $max_q|\tau_q|$. In fact, in the limit case where $\nu_q=0,\forall q$, T_λ is infinite (LTI case). Therefore, for $|\nu_q|\ll 1$ we may assume $T_\lambda\gg max_q|\tau_q|$. As a consequence, plugging (4.5.5) in (4.5.4) and using the assumption $T_\lambda\gg max_q|\tau_q|$, so that $rect_{T_\lambda}(t-\xi_\lambda-\tau_q+\tau_p)\approx rect_{T_\lambda}(t-\xi_\lambda)$, we obtain

$$\lambda^2 e^{j\phi_\lambda(t)}\simeq\sum_p\sum_q h_p^* h_q e^{-j2\pi(\nu_p-\nu_q)t}e^{-j2\pi(\nu_p-\nu_q)\tau_p}e^{j\phi_\lambda(t-(\tau_q-\tau_p))}. \tag{4.5.6}$$

Since the energy of the left-hand side, in an interval of duration T_λ, is $\lambda^2 T_\lambda$, the approximation error is smaller for the highest singular values λ. Considering also that, in view of the underspread assumption $\max_{p,q}|\tau_p\nu_q|\ll 1$, we have $e^{-j2\pi(\nu_p-\nu_q)\tau_p}\approx 1$, (4.5.6) can be approximated as

$$\lambda^2 e^{j\phi_\lambda(t)}\simeq\sum_p\sum_q h_p^* h_q e^{-j2\pi(\nu_p-\nu_q)t}e^{j\phi_\lambda(t-(\tau_q-\tau_p))}. \tag{4.5.7}$$

We are now able to prove that if $\phi_\lambda(t)$ is chosen according to (4.1.45), (4.5.7) is satisfied, at least in approximate sense, for channels with small frequency dispersion. In fact, the time-varying transfer function $H(t,f)$ of such channels is very slowly varying along t. As a consequence, the function $\phi_\lambda(t)$ resulting as a solution of (4.1.45) is also slowly varying. In fact, setting $G(t,f):=|H(t,f)|^2$, from (4.1.45) we have

$$\ddot{\phi}_\lambda(t)=-2\pi\frac{G_t(t,f_\lambda(t))}{G_f(t,f_\lambda(t))} \tag{4.5.8}$$

where $G_t(t,f):=\partial G(t,f)\partial t$ and $G_f(t,f):=\partial G(t,f)\partial f$. Thus, very small values of $G_t(t,f_\lambda(t))$ imply small values of $\ddot{\phi}_\lambda(t)$, unless $G_f(t,f_\lambda(t))$ is also very small. But $G_f(t,f_\lambda(t))$ cannot be also very small, except for a finite number of isolated points, because we assumed that the channel is dispersive in time and thus it is frequency selective. The slow variation of $\phi_\lambda(t)$ implies that we may use the following first order series expansion

$$\phi_\lambda(t-(\tau_q-\tau_p))\simeq\phi_\lambda(t)-(\tau_q-\tau_p)\dot{\phi}_\lambda(t). \tag{4.5.9}$$

Plugging this expression in (4.5.7), we obtain

$$\lambda^2 \simeq \sum_p \sum_q h_p^* h_q e^{-j2\pi(\nu_p-\nu_q)t} e^{-j(\tau_q-\tau_p)\dot{\phi}_\lambda(t)}. \tag{4.5.10}$$

Considering that

$$|H(t,f)|^2 = \sum_p \sum_q h_p^* h_q e^{-j2\pi(\nu_p-\nu_q)t} e^{-j(\tau_q-\tau_p)f} \tag{4.5.11}$$

and comparing (4.5.10) with (4.5.11), we see that the function $\phi_\lambda(t)$ that solves (4.5.10) must result as the implicit solution of the following equation

$$\lambda^2 = |H(t, \dot{\phi}_\lambda(t)/2\pi)|^2, \tag{4.5.12}$$

as in (4.1.45).

The proof in the dual case where the channel has a very small time dispersion, but a non negligible dispersion in frequency, repeats the same steps, except that the solution is sought in the frequency domain. Specifically we assume that the spectrum of $v_\lambda(t)$ has the form

$$V_\lambda(f) = \frac{1}{\sqrt{F_\lambda}} rect_{F_\lambda}(f-\zeta_\lambda) e^{j\Phi_\lambda(f)}, \tag{4.5.13}$$

and prove that an approximate solution is obtained by choosing $\Phi_\lambda(f)$ such that

$$\lambda^2 = |H(-\dot{\Phi}_\lambda(f)/2\pi, f)|^2. \tag{4.5.14}$$

Repeating the same approach seen above, applied to the kernel $\bar{h}(t,\theta)$ defined in (4.1.41), we arrive at a solution for $u_\lambda(t)$ equal to the solution for $v_\lambda(t)$.

Now we can remove the initial assumption that the channel spread function is centered around the origin of the (ν,τ) plane. In fact, recalling property ***P0***, we know that a shift of the spread function in the delay-Doppler domain induces a shift in time and a modulation of the channel output. Therefore, if we have a channel whose spread function $\bar{S}(\nu,\tau)$ is centered around (τ_0,ν_0), we may consider the corresponding channel with $S(\nu,\tau)$ centered around the origin. If $S(\nu,\tau)$ is underspread, it satisfies the properties seen before. Therefore, the channel with $\bar{S}(\nu,\tau)$ satisfies similar properties except that, for any given input, the output is shifted in the time-frequency plane. This implies for example, that if $v_\lambda(t)$ and $u_\lambda(t)$ are the right and left singular functions associated to λ, of the channel with spread function $S(\nu,\tau)$, the shifted channel with spread function $\bar{S}(\nu,\tau)$ admits as eigenfunctions the pair $v_\lambda(t)$ and $u_\lambda(t-\tau_0)e^{j2\pi\nu_0 t}$. This gives rise to the general relationship ***P3***. The same property ***P3*** states also that $v_\lambda(t)$ is a G-eigenfunction. Finally, the proof of ***P4*** is straightforward. In fact, according to (4.1.43), the eigenfunctions $w_\lambda(t)$ of a multipath channel must satisfy the following equation

$$\lambda w_\lambda(t-t_d)e^{j2\pi f_d t} = \sum_{q=0}^{Q-1} h_q w_\lambda(t-\tau_q) e^{j2\pi f_q t}. \tag{4.5.15}$$

Assuming, without any loss of generality, that $\tau_0 = 0$ and setting $t_d = 0$, $f_d = f_0$, and $\nu_q = f_q - f_0$, (4.5.15) becomes

$$\lambda w_\lambda(t) = \sum_{q=0}^{Q-1} h_q w_\lambda(t - \tau_q) e^{j2\pi\nu_q t}. \tag{4.5.16}$$

Applying the Fourier transform to both sides of (4.5.16) we get

$$\lambda W_\lambda(f) = \sum_{q=0}^{Q-1} h'_q W_\lambda(f - \nu_q) e^{-j2\pi\tau_q f}, \tag{4.5.17}$$

where $W_\lambda(f)$ is the spectrum of $w_\lambda(t)$ and $h'_q := h_q e^{j2\pi\nu_q\tau_q}$. Comparing (4.5.17) with (4.5.16), it is evident that if $max_q|\nu_q\tau_q| \ll 1$, i.e. if the channel is underspread, $h'_q \approx h_q$ and thus the eigenfunction $w_\lambda(t)$ and its spectrum $W_\lambda(f)$ must assume a similar form because they solve the same kind of equation.

4.5.2 Time-frequency representations

We recall here a few basic properties of time frequency distributions (the interested reader may refer to [18], [32], for example). Within the generalized Cohen's class of time-frequency distributions (TFD), the so called Wigner-Ville distribution (WVD) plays a prominent role because all other TFD's can be derived from the WVD through a convolution in the time-frequency domain with the smoothing function characterizing the desired distribution. Given a signal $s(t)$, its WVD is defined as

$$W_s(t, f) = \int_{-\infty}^{\infty} s^*(t - \tau/2) s(t + \tau/2) e^{-j2\pi f\tau} d\tau. \tag{4.5.18}$$

The WVD satisfies the following properties.

1.Moyal's formula: Given two signals $x(t)$ and $y(t)$, their scalar product is preserved in the time-frequency domain, in the sense that the following property, known as Moyal's formula, holds true:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_x(t, f) W_y(t, f) dt df = \left|\int_{-\infty}^{\infty} x^*(t) y(t) dt\right|^2. \tag{4.5.19}$$

2.Inversion formula: Given the WVD $W_s(t, f)$ of a signal $s(t)$, it is possible to recover $s(t)$ from $W_s(t, f)$, up to a scalar factor, using the following inversion formula:

$$s(t) = \frac{1}{s^*(0)} \int_{-\infty}^{\infty} W_s(\frac{t}{2}, f) e^{j2\pi f t} df. \tag{4.5.20}$$

3.Moments: The instantaneous frequency $f_s(t)$ of a complex signal $s(t)$ is the center of gravity of its WVD along f:

$$f_s(t) = \frac{\int_{-\infty}^{\infty} f W_s(t, f) df}{\int_{-\infty}^{\infty} W_s(t, f) df}. \tag{4.5.21}$$

By duality, the group delay $t_s(f)$ is the center of gravity of the WVD along the t:

$$t_s(f) = \frac{\int_{-\infty}^{\infty} t W_s(t, f) dt}{\int_{-\infty}^{\infty} W_s(t, f) dt}. \tag{4.5.22}$$

4.5.3 Cramér-Rao Bounds

We report here the computation of the FIM's elements. Adopting the approach based on the derivatives in the complex field proposed in [94], the parameter vector is $\boldsymbol{\xi} := (\boldsymbol{h}, \boldsymbol{\theta}, \boldsymbol{\nu}, \boldsymbol{h}^*)$ and the Fisher's information matrix (FIM) is

$$FIM(n) = \begin{pmatrix} E\{\frac{\partial f}{\partial \boldsymbol{h}}\frac{\partial f}{\partial \boldsymbol{h}^H}\} & E\{\frac{\partial f}{\partial \boldsymbol{h}}\frac{\partial f}{\partial \boldsymbol{\theta}^T}\} & E\{\frac{\partial f}{\partial \boldsymbol{h}}\frac{\partial f}{\partial \boldsymbol{\nu}^T}\} & E\{\frac{\partial f}{\partial \boldsymbol{h}}\frac{\partial f(\boldsymbol{h})}{\partial \boldsymbol{h}^T}\} \\ E\{\frac{\partial f}{\partial \boldsymbol{\theta}}\frac{\partial f}{\partial \boldsymbol{h}^H}\} & E\{\frac{\partial f}{\partial \boldsymbol{\theta}}\frac{\partial f}{\partial \boldsymbol{\theta}^T}\} & E\{\frac{\partial f}{\partial \boldsymbol{\theta}}\frac{\partial f}{\partial \boldsymbol{\nu}^T}\} & E\{\frac{\partial f}{\partial \boldsymbol{\theta}}\frac{\partial f(\boldsymbol{h})}{\partial \boldsymbol{h}^T}\} \\ E\{\frac{\partial f}{\partial \boldsymbol{\nu}}\frac{\partial f}{\partial \boldsymbol{h}^H}\} & E\{\frac{\partial f}{\partial \boldsymbol{\nu}}\frac{\partial f}{\partial \boldsymbol{\theta}^T}\} & E\{\frac{\partial f}{\partial \boldsymbol{\nu}}\frac{\partial f}{\partial \boldsymbol{\nu}^T}\} & E\{\frac{\partial f}{\partial \boldsymbol{\nu}}\frac{\partial f(\boldsymbol{h})}{\partial \boldsymbol{h}^T}\} \\ E\{\frac{\partial f}{\partial \boldsymbol{h}^*}\frac{\partial f}{\partial \boldsymbol{h}^H}\} & E\{\frac{\partial f}{\partial \boldsymbol{h}^*}\frac{\partial f}{\partial \boldsymbol{\theta}^T}\} & E\{\frac{\partial f}{\partial \boldsymbol{h}^*}\frac{\partial f}{\partial \boldsymbol{\nu}^T}\} & E\{\frac{\partial f}{\partial \boldsymbol{h}^*}\frac{\partial f(\boldsymbol{h})}{\partial \boldsymbol{h}^T}\} \end{pmatrix} \tag{4.5.23}$$

with $f(\boldsymbol{h}, \boldsymbol{\theta}, \boldsymbol{\nu})$ defined as in (4.3.2). Specifically, we have

$$\begin{aligned}
E\left\{\frac{\partial f}{\partial h_p}\frac{\partial f}{\partial h_q^*}\right\} &= \frac{1}{\sigma_v^2}\sum_{n=0}^{N-1}\sum_{k=0}^{P-1}\sum_{l=0}^{M-1}\sum_{l'=0}^{N-1} x_l(n)x_{l'}^*(n)e^{j\pi[\nu_p(nP+k+l+\theta_p)-\nu_q(nP+k+l'-\nu_q)]} \\
&\cdot\ g[\pi(k-l-\theta_p)]g[\pi(k-l'-\theta_q)] \\
E\left\{\frac{\partial f}{\partial h_p}\frac{\partial f}{\partial h_q}\right\} &= 0 \\
E\left\{\frac{\partial f}{\partial h_p}\frac{\partial f}{\partial \theta_q}\right\} &= \frac{\pi h_q^*}{\sigma_v^2}\sum_{n=0}^{N-1}\sum_{k=0}^{P-1}\sum_{l=0}^{M-1}\sum_{l'=0}^{N-1} x_l(n)x_{l'}^*(n)e^{j\pi[\nu_p(nP+k+l+\theta_p)-\nu_q(nP+k+l'-\nu_q)]} \\
&\cdot\ g[\pi(k-l-\theta_p)]\{-j\nu_q g[\pi(k-l'-\theta_q)] - \dot{g}[\pi(k-l'-\theta_q)]\} \\
E\left\{\frac{\partial f}{\partial h_p}\frac{\partial f}{\partial \nu_q}\right\} &= \frac{-j\pi h_q^*}{\sigma_v^2}\sum_{n=0}^{N-1}\sum_{k=0}^{P-1}\sum_{l=0}^{M-1}\sum_{l'=0}^{N-1} x_l(n)x_{l'}^*(n)e^{j\pi[\nu_p(nP+k+l+\theta_p)-\nu_q(nP+k+l'-\nu_q)]} \\
&\cdot\ (nP+k+l'+\theta_q)g[\pi(k-l-\theta_p)]g[\pi(k-l'-\theta_q)] \\
E\left\{\frac{\partial f}{\partial \theta_p}\frac{\partial f}{\partial \theta_q}\right\} &= \frac{2\pi^2}{\sigma_v^2}\mathcal{R}e\{h_p h_q^*\sum_{n=0}^{N-1}\sum_{k=0}^{P-1}\sum_{l=0}^{M-1}\sum_{l'=0}^{N-1} e^{j\pi[\nu_p(nP+k+l+\theta_p)-\nu_q(nP+k+l'-\theta_q)]} \\
&\cdot\ [j\nu_p g(\pi(k-l-\theta_p)) - \dot{g}(\pi(k-l-\theta_p))]x_l(n)x_{l'}^*(n) \\
&\cdot\ [-j\nu_q g(\pi(k-l'-\theta_q)) - \dot{g}(\pi(k-l'-\theta_q))]\} \\
E\left\{\frac{\partial f}{\partial \theta_p}\frac{\partial f}{\partial \nu_q}\right\} &= \frac{2\pi^2}{\sigma_v^2}\mathcal{R}e\{h_p h_q^*\sum_{n=0}^{N-1}\sum_{k=0}^{P-1}\sum_{l=0}^{M-1}\sum_{l'=0}^{N-1} x_l(n)x_{l'}^*(n)e^{j\pi[\nu_p(nP+k+l+\theta_p)-\nu_q(nP+k+l'-\theta_q)]} \\
&\cdot\ [\nu_p g(\pi(k-l-\theta_p)) + j\dot{g}(\pi(k-l-\theta_p))](nP+k+l'+\theta_q)g(\pi(k-l'-\theta_q))\} \\
E\left\{\frac{\partial f}{\partial \nu_p}\frac{\partial f}{\partial \nu_q}\right\} &= \frac{2\pi^2}{\sigma_v^2}\mathcal{R}e\{h_p h_q^*\sum_{n=0}^{N-1}\sum_{k=0}^{P-1}\sum_{l=0}^{M-1}\sum_{l'=0}^{N-1} x_l(n)x_{l'}^*(n)e^{j\pi[\nu_p(nP+k+l+\theta_p)-\nu_q(nP+k+l'-\theta_q)]} \\
&\cdot\ (nP+k+l+\theta_p)(nP+k+l'+\theta_q)g[\pi(k-l-\theta_p)]g[\pi(k-l'-\theta_q)]\}
\end{aligned}$$

BIBLIOGRAPHY

[1] N. Al-Dhahir, "Time-varying versus time-invariant finite-length MMSE-DFE on stationary dispersive channels", *IEEE Transactions on Communications*, Vol. 46, No. 1, pp. 11 –15, Jan. 1998.

[2] N. Al-Dhahir, J.M. Cioffi, "Block Transmission over Dispersive Channels: Transmit Filter Optimization and Realization, and MMSE-DFE Receiver Performance," *IEEE Trans. Inform. Theory*, pp. 137-160, Jan. 1996.

[3] M. Alles, S. Pasupathy, "Channel knowledge and optimal performance for two-wave Rayleigh fading channels", *IEEE Transactions on Vehicular Technology*, Vol. 43, No. 1, pp.8 –20, Feb. 1994.

[4] M.S. Alouini A.J. Goldsmith, "Adaptive modulation for Nakagami fading channels", *Proc. of IEEE GLOBECOM' 97*, Phoenix, November 1997.

[5] M.-S. Alouini, X. Tang, A. Goldsmith, "An adaptive modulation scheme for simultaneous voice and data transmission over fading channels", *Proc. of 48th IEEE Vehicular Technology Conference, VTC-98*, vol.2, pp. 939 - 943, 18-21 May 1998.

[6] F. Auger and P. Flandrin, "Improving the Readability of Time-Frequency and Time-Scale Representations by the Reassignment Method", *IEEE Trans. on Signal Proc.*, Vol.43, No.5, May 1995, pp.1068-1089.

[7] E. Baccarelli, R. Cusani, S. Galli, "A novel adaptive receiver with enhanced channel tracking capability for TDMA-based mobile radio communications", *IEEE Journal on Selected Areas in Communications*, Vol. 16, No. 9, pp. 1630–1639, Dec. 1998.

[8] S. Barbarossa, A. Scaglione,"On the capacity of linear time-varying channels", *Proc. of 1999 IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, ICASSP-99*, pp. 2627–2630, Phoenix, March 15-19, 1999.

[9] S. Barbarossa, A. Scaglione, "Optimal precoding for transmissions over linear time-varying channels", *Proc. of IEEE Global Telecommunications Conference, GLOBECOM '99*, Rio de Janeiro, Dec. 1999.

[10] P.A. Bello,"Characterization of randomly time-variant linear channels", *IEEE Trans. on Circuits and Systems*, vol. CS-11, No. 4, pp. 360–393, Dec. 1963.

[11] S. Bhashyam, A.M. Sayeed, B. Aazhang, "Time-selective signaling and reception for multipath fading channels", *Proc. of 1998 IEEE International Symposium on Information Theory*, p. 157, 16-21 Aug. 1998.

[12] D.K. Borah, B.D. Hart, "A robust receiver structure for time-varying, frequency-flat, Rayleigh fading channels", *IEEE Transactions on Communications*, Vol. 47, No. 3, pp. 360–364, March 1999.

[13] G.E. Bottomley, S. Chennakeshu, "Unification of MLSE receivers and extension to time-varying channels" *IEEE Transactions on Communications*, Vol. 46, No. 4, pp. 464–472, April 1998.

[14] G. Caire, S. Shamai, "On the capacity of some channels with channel state information", *Proc. 1998 IEEE International Symposium on Information Theory*, p. 42, 16-21 Aug. 1998.

[15] G. Caire, G. Taricco, E. Biglieri, "Optimum power control over fading channels", *IEEE Trans. Inform. Theory*, Vol.45, No. 5, pp. 1468–1489, July 1999.

[16] J.K. Cavers, "Variable-rate transmission for Rayleigh fading channels", *IEEE Trans. on Comm.*, Vol.20, pp.15-22, February 1972.

[17] B.-S. Chen, Y.-C. Chung, D.-F. Huang, "Optimal time-frequency deconvolution filter design for nonstationary signal transmission through a fading channel: A filter bank approach", *IEEE Transactions on Signal Processing*, Vol. 46, No. 12, pp. 3220 –3234, Dec. 1998.

[18] L. Cohen, "Time-Frequency Analysis", Prentice-Hall, Englewood Cliffs, New Jersey, 1995.

[19] P.M. Crespo, J. Jimenez, "Computer simulation of radio channels using a harmonic decomposition technique", *IEEE Transactions on Vehicular Technology*, Vol. 44, No. 3, pp. 414 – 419, Aug. 1995.

[20] I. Daubechies, "Time-frequency localization operators: A geometric phase-space approach", *IEEE Trans. Info. Theory*, Vol. 34, pp. 605–612, July 1988.

[21] L.M. Davis, I.B. Collings, R.J. Evans, "Coupled estimators for equalization of fast-fading mobile channels", *IEEE Transactions on Communications*, Vol. 46, No. 10, pp. 1262– 1265, Oct. 1998.

[22] L.M. Davis, I.B. Collings, R.J. Evans, "Constrained maximum likelihood estimation of time-varying linear channels", *Proc. of First IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications*, pp. 1 - 4, Paris, 16-18 April 1997.

[23] S.N. Diggavi, A. Paulraj, "Signal detection for time-varying vector channels", *Conf. Record of the Twenty-Ninth Asilomar Conference on Signals, Systems and Computers*, vol.1, pp. 152 – 156, 30 Oct.-2 Nov. 1995.

[24] S.N. Diggavi, "Analysis of multicarrier transmission in time-varying channels", *Proc. of 1997 IEEE Int. Conference on Communications, ICC '97*, vol.3, pp.1191 – 1195 , Montreal, 8-12 June 1997.

[25] L. Dossi, G. Tartara, F. Tallone, "Statistical analysis of measured impulse response functions of 2.0 GHz indoor radio channels", *IEEE Journal on Selected Areas in Communications*, Vol. 14, No. 3, pp. 405 – 410, April 1996.

[26] M. Effros, "Robustness to channel variation in source coding for transmission across noisy channels", *Proc. of 1997 IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, ICASSP-97*, vol.4, pp. 2961 - 2964, 21-24 April 1997.

[27] European Telecommunications Standard, "Radio broadcast systems for television, sound and data services: Framing structure, channel coding and modulation for digital terrestrial television," ETS 300 744, March 1997.

[28] ETSI Normalization Commitee, Norme ETSI, doc. 3ERI085B, Sophia-Antipolis, Valbonne, France 1998.

[29] T. Eyceoz, A. Duel-Hallen, H. Hallen, "Deterministic channel modeling and long range prediction of fast fading mobile radio channels", *IEEE Communications Letters*, Vol. 2, No. 9, pp. 254 – 256, Sept. 1998.

[30] A. Farina, A. Russo, F. Scannapieco, S. Barbarossa, "Theory of radar detection in coherent Weibull clutter", *IEE Proceedings*, Vol. 134, Pt. F, pp. 174–190, April 1987.

[31] P. Flandrin, "Maximum signal energy concentration in the time-frequency plane", IEEE ICASSP '88, pp. 2176–2179, New York 1988.

[32] P. Flandrin, J. Stockler (Translator), "Time-Frequency/Time-Scale Analysis (Wavelet Analysis and Its Applications, V. 10)", Academic Press, 1999.

[33] R. Gallager, *Elements of Information Theory*, Section 8, New York: Wiley, 1968.

[34] G.B. Giannakis, C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels" *Proc. of the IEEE*, Vol. 86, No. 10 , pp. 1969–1986, Oct. 1998.

[35] G.B. Giannakis, C. Tepedelenlioglu, H. Liu, "Adaptive blind equalization of time-varying channels", *Proc. of 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-97*, vol.5, pp. 4033 – 4036, 21-24 April 1997.

[36] G. B. Giannakis, "Blind equalization of time-varying channels: a deterministic multichannel approach", *Proc. of the 8th IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, pp. 180 – 183, Corfu, 24-26 June 1996.

[37] A. Goldsmith, "Design and performance of high-speed communication systems over time-varying radio channels", *PhD thesis*, Univ. of California, Berkeley, Sept. 1994.

[38] A. Goldsmith, "Multiuser capacity of cellular time-varying channels", *Conf. Record of the Twenty-Eighth Asilomar Conference on Signals, Systems and Computers*, vol.1, pp. 83 – 88, 31 Oct.-2 Nov. 1994.

[39] A. Goldsmith, "Joint source/channel coding for wireless channels", *Proc. of 1995 IEEE 45th Vehicular Technology Conference, VTC-95*, vol.2, pp. 614 – 618, 25-28 July 1995.

[40] A.J. Goldsmith, P.P. Varaiya, "Capacity, mutual information, and coding for finite-state Markov channels", *IEEE Transactions on Information Theory*, Vol. 42, No. 3, pp. 868 – 886, May 1996.

[41] A. J. Goldsmith, S. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.,* vol. 45, pp. 1218-1230, Oct. 1997.

[42] A. J. Goldsmith, P. P. Varaiya, "Capacity of Fading Channels with Channel Side Information," *IEEE Trans. on Info. Theory,* pp. 1986-1992, Nov. 1997.

[43] D.L. Goeckel, "Adaptive coding for time-varying channels using outdated fading estimates", *IEEE Trans. on Comm.*, Vol.47, pp.844-855, June 1999.

[44] D. J. Grenwood, L. Hanzo, "Characterization of Mobile Radio Channels", in *Mobile Radio Communications*, ed. by R. Steele, *IEEE Press*, Piscataway, NJ, 1992.

[45] W.C. Jakes, *Microwave Mobile Communications*, Section 1.5, IEEE Press, Piscataway, NJ, 1974.

[46] B.D. Hart, D.P. Taylor, "Maximum-likelihood synchronization, equalization, and sequence estimation for unknown time-varying frequency-selective Rician channels", *IEEE Transactions on Communications*, Vol. 46, No. 2, pp. 211–221, Feb. 1998.

[47] P. Hoeher, "A statistical discrete-time model for the WSSUS multipath channel", *IEEE Trans. on Vehicular Technology*, vol. 41, pp. 461-467, Nov. 1992.

[48] S. Kasturia, J. T. Aslanis, and J. M. Cioffi, "Vector Coding for Partial Response Channels," *IEEE Trans. on Info. Theory,* vol. 36, pp. 741-761, July 1990.

[49] M. Khansari, M. Vetterli, "Source coding and transmission of signals over time-varying channels with side information", *Proc. of the 1995 IEEE International Symposium on Information Theory*, pp. 140, 17-22 Sept. 1995.

[50] W. Kozek "On the transfer function calculus for underspread LTV channels" *IEEE Transactions on Signal Processing*, Vol. 45, No. 1, pp. 219–223, Jan. 1997.

[51] W. Kozek, A.F. Molisch, "On the eigenstructure of underspread WSSUS channels", *First IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications*, pp. 325 – 328, Paris, 16-18 April 1997.

[52] I. Kozintsev, K. Ramchandran, "Multiresolution joint source-channel coding using embedded constellations for power-constrained time-varying channels", *Proc. of the 1996 IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, ICASSP-96*, vol. 4, pp. 2343 - 2346, Atlanta, 7-10 May 1996.

[53] H. Kubo, K. Murakami, T. Fujino, "An adaptive maximum-likelihood sequence estimator for fast time-varying intersymbol interference channels" *IEEE Transactions on Communications*, Vol. 42, No. 2, pp. 1872 - 1880, FEBRUARY/MARCH/APRIL 1994.

[54] T. Kurner, D.J. Cichon, W. Wiesbeck, "Evaluation and verification of the VHF/UHF propagation channel based on a 3-D-wave propagation model", *IEEE Transactions on Antennas and Propagation*, Vol. 44, No.3, pp. 393 – 404, March 1996.

[55] H. J. Landau, H. Pollack, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty - II", *The Bell system Techn. Journal*, pp. 65–84, Jan. 1961.

[56] H.A. Leinhos, "Capacity calculations for rapidly fading communications channels", *IEEE Journal of Oceanic Engineering*, Vol. 21, No. 2, pp. 137 - 142, April 1996.

[57] W. S. Leon, D.P. Taylor, "An adaptive receiver for the time- and frequency-selective fading channel", *IEEE Transactions on Communications*, pp. 1548 - 1555, Vol. 45, No. 12, Dec. 1997.

[58] Y. Li, L.J. Cimini, N.R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels", *IEEE Transactions on Communications*, Vol. 46, No. 7, pp. 902 – 915, July 1998.

[59] J. Lin, J.G. Proakis, L. Fuyun, H. Lev-Ari, "Optimal tracking of time-varying channels: a frequency domain approach for known and new algorithms", *IEEE Journal on Selected Areas in Communications*, Vol. 13, No. 1 , pp. 141 -154, Jan. 1995.

[60] H. Liu, G.B. Giannakis, "Deterministic approaches for blind equalization of time-varying channels with antenna arrays", *IEEE Transactions on Signal Processing*, Vol. 46, No. 11, pp. 3003–3013, Nov. 1998.

[61] G. Matz, F. Hlawatsch "Time-Frequency Transfer Function Calculus (Symbolic Calculus) of Linear Time-Varying Systems (Linear Operators) Based on a Generalized Underspread Theory", *Journal of Mathematics Physics* (Special issue on Wavelets and Time-Frequency Analysis), vol. 39, no. 8, pp.4041-4070, Aug. 1998.

[62] M. Medard, R.G. Gallager, "The effect of a randomly time-varying channel on mutual information", *Proc. of the 1995 IEEE International Symposium on Information Theory*, p. 139, 17-22 Sept. 1995.

[63] M. Medard, "The capacity of time-varying multiple user channels in wireless communications", PhD Thesis, MIT, Boston, Sept. 1995.

[64] M. Medard, R.G. Gallager, "The issue of spreading in multipath time-varying channels", *Proc. of 1995 IEEE 45th Vehicular Technology Conf.*, vol.1, pp. 1 - 5, 25-28 July 1995.

[65] M. Medard, R.G. Gallager, "The effect of channel variations upon capacity", *Proc. of 1996 IEEE 46th Vehicular Technology Conference*, vol.3, pp. 1781 - 1785, 28 April-1 May 1996.

[66] M. Medard, A. Goldsmith, "Capacity of time-varying channels with channel side information", *Proc. of 1997 IEEE Int. Symposium on Information Theory*, p. 372, 29 June-4 July 1997.

[67] A. Narula, M.J. Lopez, M.D. Trott, G.W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels", *IEEE Journal on Selected Areas in Communications*, Vol. 16, No. 8, pp. 1423–1436, Oct. 1998.

[68] B.C. Ng, S.N. Diggavi, A. Paulraj, "Joint structured channel and data estimation over time-varying channels", *Proc. of IEEE Global Telecommunications Conference, GLOBECOM '97*, vol.1, pp. 409 - 413, 3-8 Nov. 1997.

[69] A. Papoulis, *Signal Analysis*, New York: McGraw-Hill International Ed., 1977.

[70] D. Parsons, *The Mobile Radio Propagation Channel*, Pentech Press, Publ. London, 1992.

[71] J. G. Proakis, *Digital Communications*, (second ed.), New York: McGraw-Hill International Ed., 1989.

[72] G.G. Raleigh, A. Paulraj, "Time varying vector channel estimation for adaptive spatial equalization", *Proc. of IEEE Global Telecommunications Conference, GLOBECOM '95*, vol.1, pp. 218 – 224, 13-17 Nov. 1995.

[73] G.G. Raleigh, T. Boros, "Joint space-time parameter estimation for wireless communication channels", *IEEE Transactions on Signal Processing*, Vol. 46, No. 5, pp. 1333 - 1343, May 1998.

[74] T.S. Rappaport, S. Sandhu, "Radio-wave propagation for emerging wireless personal-communication systems", *IEEE Antennas and Propagation Magazine*, Vol. 36, No. 5, pp. 14 – 24, Oct. 1994.

[75] M. Reinhardt, M. Schurr, J. Lindner, "Double-equalization-algorithm for time and frequency selective multiuser channels", *Proc. of IEEE 4th International Symposium on Spread Spectrum Techniques and Applications*, vol.2, pp. 831 – 836, 22-25 Sept. 1996.

[76] W.E. Ryan, L. Han, "Modulation and coding for short-duration deterministically time-varying channels", *Proc. of 1995 IEEE Int. Conf. on Communications, ICC '95*, vol.3, pp. 1790 – 1794, Seattle, 18-22 June 1995.

[77] A.M. Sayeed, B. Aazhang, "Joint multipath-Doppler diversity in mobile wireless communications", *IEEE Transactions on Communications*, Vol. 47, No. 1,pp. 123 – 132, Jan. 1999.

[78] A. Scaglione, S. Barbarossa and G. B. Giannakis, "Direct Equalization of Time-Selective Fading Channels Using Redundant Filterbank Precoders", *Proc. of the IEEE Conf. on Digital Signal Proc.*, Bryce Canyon (UT), August 1998.

[79] A. Scaglione, S. Barbarossa, and G. B. Giannakis, "Filterbank Transceivers Optimizing Information Rate in Block Transmissions over Dispersive Channels," *IEEE Trans. on Info. Theory*, Vol. 45, No. 3, pp. 1019 -1032, April 1999.

[80] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant Filterbank Precoders and Equalizers - Part I: Unification and Optimal Designs," *IEEE Trans. on Signal Processing*, Vol. 47, No. 7, pp. 1988–2006, July 1999.

[81] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant Filterbank Precoders and Equalizers - Part II: Blind Channel Estimation, Synchronization, and Direct Equalization," *IEEE Trans. on Signal Processing*, Vol. 47, No. 7, pp. 2007–2022, July 1999.

[82] B. Senadji, A.J. Levy, "A statistical model for the simulation of time-varying multipath mobile radio propagation channel", *Proc. of 1994 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-94*, vol.6, pp. VI/149 - VI/152, 19-22 April 1994.

[83] N. Seshadri, "Joint data and channel estimation using blind trellis search techniques", *IEEE Transactions on Communications*, pp. 1000 - 1011, February-April 1994.

[84] N. Sheikholeslami, P. Kabal, "Linear time varying precoder applied to an ISI channel", *Proc. of 1997 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, vol.1, pp. 36 – 39, 20-22 Aug. 1997.

[85] L. Sirovich, B. W. Knight, "On the eigentheory of operators which exhibit a slow variation", *Quart. Appl. Math.*, Vol. 38, pp. 469-488, 1980.

[86] G.E. Shilov, *Elementary Real and Complex Analysis*, Dover Publications, Inc., New York, 1973, p. 252.

[87] D. Slepian, H. Pollack, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty - I", *The Bell system Techn. Journal*, pp. 43–64, Jan. 1961.

[88] P. Stoica, R. L. Moses, *Introduction to Spectral Analysis*, Prentice Hall, 1997.

[89] Suzuki, "A statistical model for urban radio propagation", *IEEE Transactions on Communications*, Vol. 25, pp. 673–680, 1977.

[90] C. Tepedelenlioğlu, "Deterministic blind estimation of time- and frequency selective fading channels using filterbank precoders", *Proc. of the 2nd IEEE Workshop on Signal Proc. Advances in Wireless Communications*, pp. 74 – 77, Annapolis (MD), May 9-12, 1999.

[91] M.K. Tsatsanis, G.B. Giannakis, "Equalization of rapidly fading channels: self-recovering methods", *IEEE Transactions on Communications*, Vol. 44, No. 5, pp. 619–630, May 1996.

[92] M.K. Tsatsanis, G.B. Giannakis, G. Zhou, "Estimation and equalization of fading channels with random coefficients", *Signal Processing*, vol. 53, No. 2-3, pp.211-229, Sept. 1996.

[93] M.K. Tsatsanis, G.B. Giannakis, "Subspace methods for blind estimation of time-varying FIR channels", *IEEE Transactions on Signal Processing*, Vol. 45, No. 12 , pp. 3084–3093, Dec. 1997.

[94] A. van den Bos, "A Cramér-Rao lower bound for complex parameters", *IEEE Trans. on Signal Proc.*, Vol. 42, p.2859, Oct. 1994.

[95] T. Wang, V.K. Dubey, "Generation of scattering functions by computer simulation for mobile communication channels", *Proc. of the IEEE 46th Vehicular Technology Conference, VTC-96*, vol.3, pp. 1443 - 1447, 28 April-1 May 1996.

[96] W.T. Webb, R. Steele, "Variable rate QAM for mobile radio", *IEEE Trans. on Comm.*, Vol.43, pp.2223-2230, July 1995.

[97] B.D. Woerner, J.H. Reed, T.S. Rappaport, "Simulation issues for future wireless modems", *IEEE Communications Magazine*, Vol. 32, No. 7, pp. 42– 53, July 1994.

[98] C. Y. Wong, R. S. Cheng, K. B. Letaief, R. D. Ross,"Multiuser OFDM with adaptive subcarrier, bit, and power allocation", *IEEE Trans. on Sel. Areas in Comm.*, Vol.17, pp.1747-1757, October 1999.

[99] K. Yao, "A representation theorem and its applications to spherically-invariant random processes", *IEEE Trans. on Inform. Theory*, pp. 600–608, 1973.

[100] K.W. Yip, T.-S. Ng, "Karhunen-Loeve expansion of the WSSUS channel output and its application to efficient simulation", *IEEE Journal on Selected Areas in Communications*, Vol. 15, No. 4, pp. 640 – 646, May 1997.