Principles of Channel Estimation in OFDM

Half-day Tutorial

Siemens AG
Com MN PG R S C4
29. April 2005

Technische Universität München
Univ.-Prof. Dr.-Ing. Wolfgang Utschick

Contents

Notation

Vectors and Matrices

$$\boldsymbol{x} \in \mathbb{C}^N \quad \Leftrightarrow \quad \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} = [x_1, \ldots, x_N]^{\mathsf{T}}$$

$$\boldsymbol{X} \in \mathbb{C}^{M \times N} \quad \Leftrightarrow \quad \begin{bmatrix} x_{1,1} & \cdots & x_{1,N} \\ \vdots & & \vdots \\ x_{M,1} & \ldots & x_{M,N} \end{bmatrix}$$

Random Variables and Realizations

$$\mathsf{x} \sim f_{\mathsf{x}}(\xi) \quad \Leftrightarrow \quad F_{\mathsf{x}}(a) = P_{\mathsf{x}}(\xi \leq a)$$

$$\mathsf{E}[\mathsf{x}] = \int_{\mathbb{R}} \mathsf{x} \, \mathsf{d} \, P_{\mathsf{x}} = \int_{\mathbb{R}} \xi \; f_{\mathsf{x}}(\xi) \, \mathsf{d} \, \xi$$

Principles of OFDM

1. Transmission over Orthogonal Basis Functions

Orthonormal systems (ONSs) are ideal for data transmission by means of multiple access techniques.

In ONSs normalized orthogonal basis functions[1] $\psi_q(t)$ serve as independent carriers of information for multiple data streams.

The maximum number of independent data streams is equal to the dimension N of the given signal vector space.

For band-limited transmission systems with finite access time per channel use the dimension of the signal vector space is finite:

$$\boxed{N = BT,} \tag{1.1}$$

with bandwidth $B = 2f_\mathrm{G}$ and channel access time T.[2]

[1] An ONS consists of a set of basis functions $\{\psi_i(t)\}$ which are characterized by $\langle\psi_i, \psi_j\rangle = \delta_{i,j}$ for all i and j.

[2] N equals the multiple of the channel access time to the required sampling intervall $\frac{1}{B}$.

1.1 Dimension, Modulation, Demodulation

Given the set of time instants ($\rightarrow$ b-th channel use[3])

$$\mathbb{T}_b = \{t \mid bT \leq t < (b+1)T\} \quad \subset \quad \mathbb{R}, \tag{1.2}$$

the data transmission is based on an orthonormal system of basis functions (Fig. 1.1) from the Hilbert space

$$\mathcal{L}_2(\mathbb{T}_b) = \{x(t) \mid \|x\|_2 < \infty,\ t \in \mathbb{T}_b \ \wedge\ x(t) = 0, t \notin \mathbb{T}_b\}\,. \tag{1.3}$$

The modulation at the receiver is equal to

$$\boxed{x_b(t) = \sum_{q=1}^{N} \underbrace{x_{b,q}}_{\text{Data}} \overbrace{\psi_{b,q}(t)}^{\text{Basis}},} \tag{1.4}$$

with $x_{b,q} \sim \mathsf{x}_{b,q}$ as the realization of the q-th stochastic random data process and the orthonormal basis functions

$$\psi_{b,1}(t), \psi_{b,2}(t), \ldots, \psi_{b,N}(t) \quad \in \quad \mathcal{L}_2(\mathbb{T}_b). \tag{1.5}$$

N equals the dimension of the transmission system.

[3] A generalized definition of $\mathbb{T}_b$ reads $\mathbb{T}_b = \{t \mid -T_\mathsf{p} + b(T+T_\mathsf{p}) \leq t < b(T+T_\mathsf{p}) + T\}$.

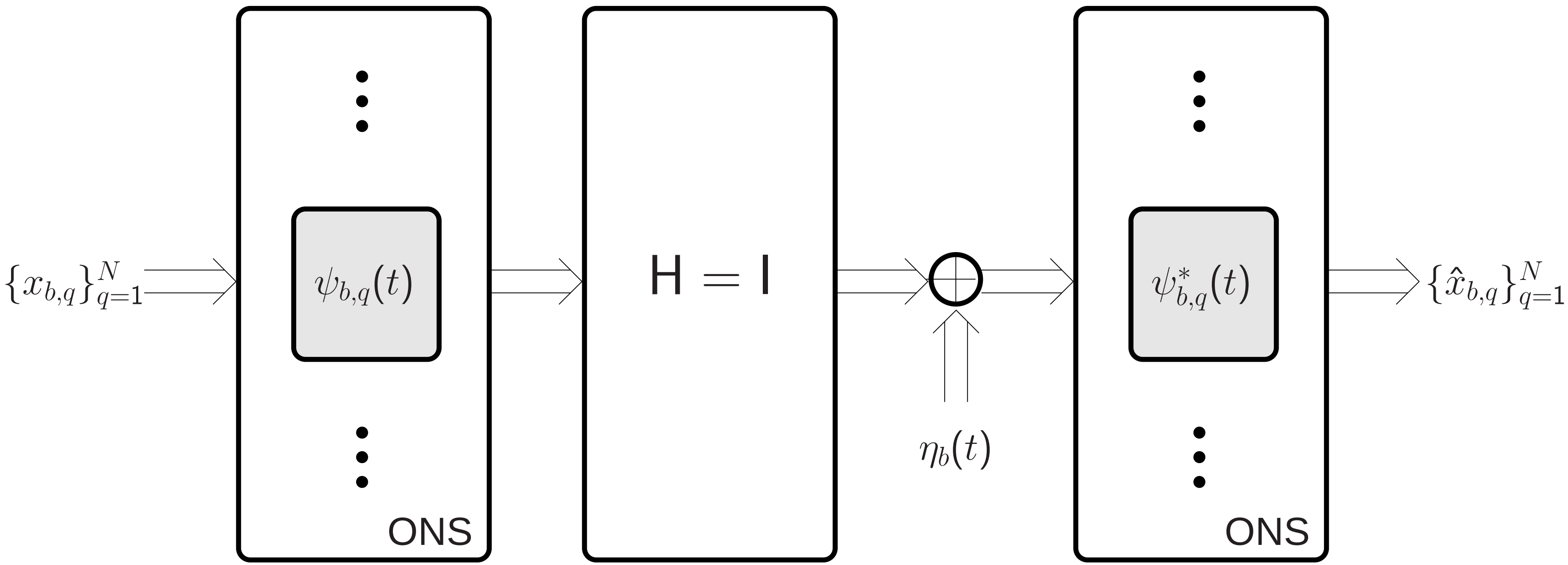

Fig. 1.1. Data Transmission by means of an Orthonormal System of Basis Functions

1.2 Additive White Gaussian Noise Channel

Given an AWGN transmission channel $\mathsf{H} : \mathcal{L}_2(\mathbb{T}_b) \rightarrow \mathcal{L}_2(\mathbb{T}_b)$ ($\rightarrow$ identity operator[4]), the signal model at the receiver reads

$$\begin{aligned} y_b(t) &= (\mathsf{H}x_b)(t) + \eta_b(t) \quad (1.6) \\ &= \sum_{q=1}^{N} x_{b,q}\, \psi_{b,q}(t) + \eta_b(t), \end{aligned}$$

with $\eta_b(t)$ als sample function of the random noise process at the receiver. The demodulation of the $x_{b,q}$ is again based on the orthonormal basis functions (Fig. 1.1)

$$\psi_{b,1}(t), \psi_{b,2}(t), \ldots, \psi_{b,N}(t) \quad \in \quad \mathcal{L}_2(\mathbb{T}_b), \tag{1.7}$$

i. e.

$$\boxed{\hat{x}_{b,q} = \langle y, \psi_{b,q}\rangle + \langle \eta_b, \psi_{b,q}\rangle = x_{b,q} + \eta_{b,q}, \qquad \forall q = 1, \ldots, N.} \tag{1.8}$$

[4]For any orthonormal system $\{\psi_{b,q}(t)\}_{q=1}^{Q}$ the expansion of the identity operators is given by $\boldsymbol{I} = \sum_{q=1}^{N} \langle \bullet, \psi_{b,q}\rangle \psi_{b,q}(t)$.

1.3 Orthonormal Systems for Non-Trivial Channels

Since the transmission channel generally cannot be described by the idendity operator, the basis functions $\{\psi_{b,q}(t)\}_{q=1}^{N}$ and $\{\varphi_{b,q}(t)\}_{q=1}^{N}$[5] from the series expansion

$$\boxed{\mathsf{H} = \textstyle\sum_{q=1}^{N} \gamma_q \, \varphi_{b,q}(t) \langle \bullet, \psi_{b,q} \rangle,} \tag{1.9}$$

will not be identical with any prior set of basis functions for data transmission. In order to avoid interference, the basis functions for data transmission must be adopted from (1.9). Finally, the following signal model at the receiver can be obtained as (Fig. 1.2)

$$\begin{aligned} y_b(t) \; &= \; (\mathsf{H} x_b)(t) + \eta_b(t) = \sum_{q=1}^{N} \gamma_q \langle x_b, \psi_{b,q} \rangle \varphi_{b,q}(t) + \eta_b(t) \qquad (1.10) \\ &= \; \sum_{q=1}^{N} \gamma_q \left\langle \sum_{q'=1}^{N} x_{b,q'} \, \psi_{b,q'}, \psi_{b,q} \right\rangle \varphi_{b,q}(t) + \eta_b(t) = \sum_{q=1}^{N} \gamma_q x_{b,q} \, \varphi_{b,q}(t) + \eta_b(t). \end{aligned}$$

[5] Given a compact operators A, its singular value decomposition equals: $\mathsf{A} = \sum_{q=1}^{\infty} \sigma_q \langle \bullet, \psi_q \rangle \varphi_q(t)$. The orthonormal basis $\{\psi_q(t)\}_{q=1}^{\infty}$ and $\{\varphi_q(t)\}_{q=1}^{\infty}$ depend on operator A.

Demodulation of data $x_{b,q}$ is performed by means of the orthonormal basis function (Fig. 1.2)

$$\varphi_{b,1}(t), \varphi_{b,2}(t), \ldots, \varphi_{b,N}(t) \quad \in \quad \mathcal{L}_2(\mathbb{T}_b), \tag{1.11}$$

i. e.

$$\boxed{\hat{x}_{b,q} = \langle y, \varphi_{b,q} \rangle + \langle \eta, \varphi_{b,q} \rangle = H_q x_{b,q} + \eta_{b,q}, \qquad \forall q = 1, \ldots, N.} \tag{1.12}$$

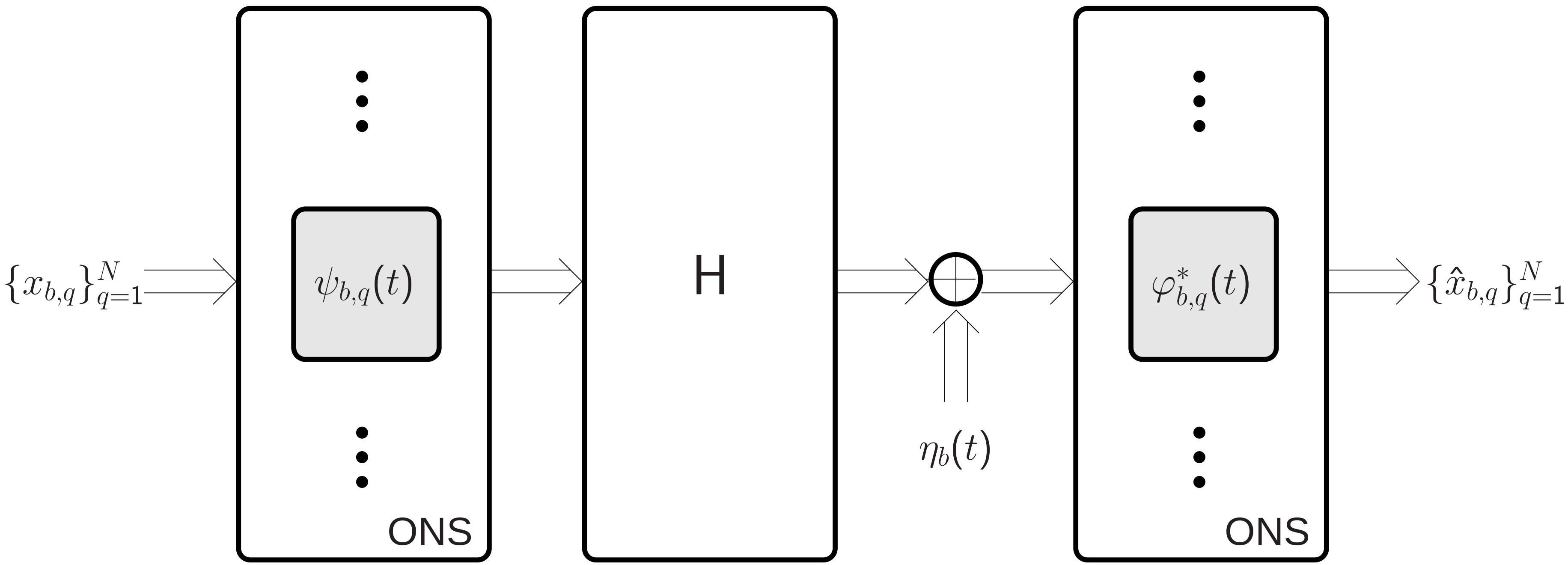

Fig. 1.2. Data Transmission over Channel-depending Basis Functions

1.4 Maximum Information Rate (Capacity)

Given a stationary gaussian modulation and gaussian distributed noise processes, i. e.

$$x_{b,q} \sim \mathsf{N}\left(0, \sigma^2_{x,q}\right) \quad \text{und} \quad \eta_{b,q} \sim \mathsf{N}\left(0, \sigma^2_{\eta,q}\right), \tag{1.13}$$

the mutual information is

$$I(x,y) = \frac{B}{N} \sum_{q=1}^{N} \log_2 \left(1 + \frac{\sigma^2_{x,q}}{\sigma^2_{\eta,q}} \gamma_q^2 \right). \tag{1.14}$$

The maximum information rate (capacity) for a given transmit power E_{Tx}, i. e. $\sum_{q=1}^{N} \sigma_{x,q} \leq E_{\mathsf{Tx}}$, can be achieved by the distribution of E_{Tx} over N basis functions with respect to the *Waterfilling*-principle:

$$\sigma^2_{x,q} = \max\left(\mu - \frac{\sigma^2_{\eta,q}}{\gamma_q^2}, 0\right), \quad \text{and} \quad \mu = \frac{1}{N}\left(E_{\mathsf{Tx}} + \sum_{q=1}^{N} \frac{\sigma^2_{\eta,q}}{\gamma_q^2} \chi_q \right), \tag{1.15}$$

where $\chi_q = \begin{cases} 1 \;;\; \sigma^2_{x,q} > 0 \\ 0 \;;\; \text{sonst} \end{cases}$.[6]

[6]The $\mu > 0$ and the optimum power distribution $\{\sigma^2_{x,q}\}_{q=1}^{N}$ is computed iteratively.

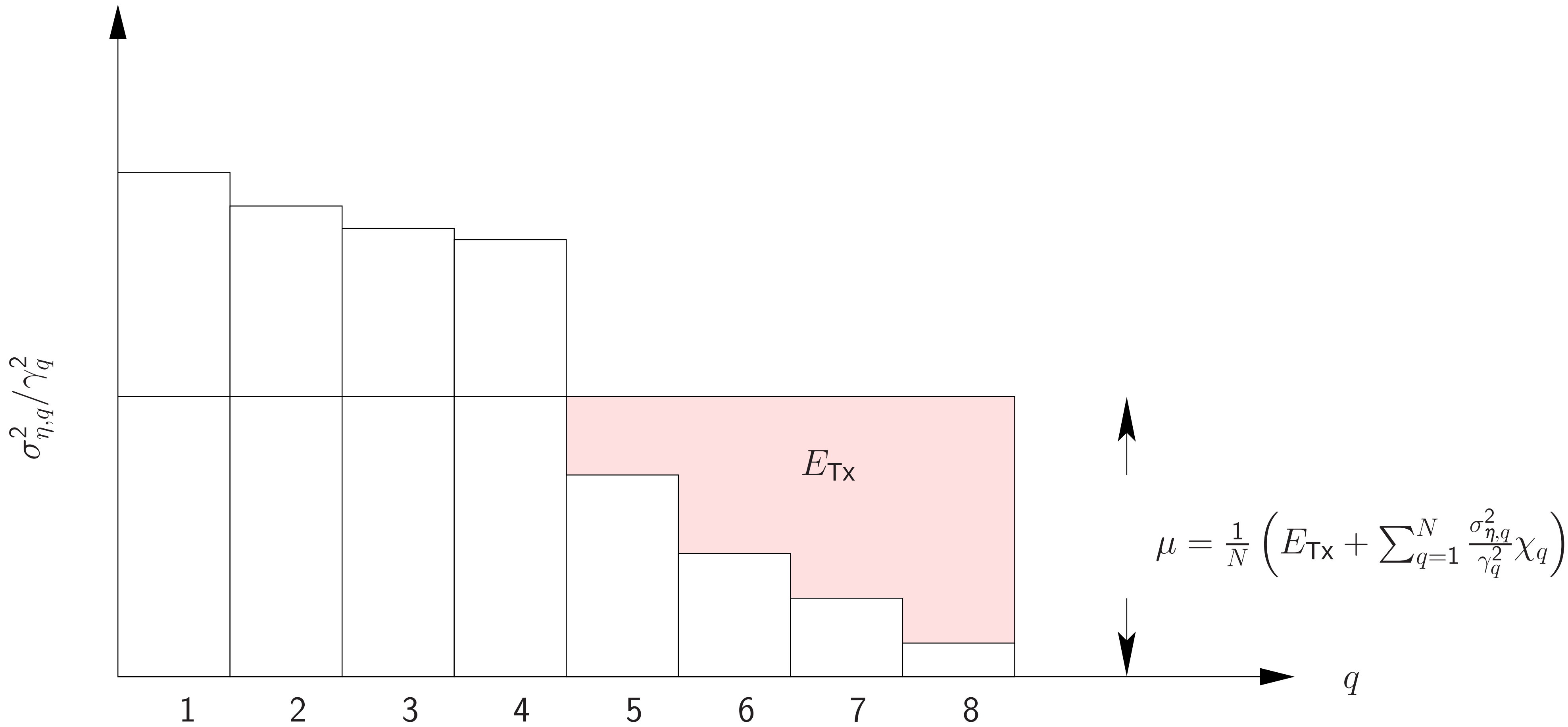

Fig. 1.3. Waterfilling Principle

2. Orthogonal Frequeny Division Multiplexing (OFDM)

Choosing

$$\psi_{b,q}(t) = p_{\mathbb{T}_b}(t) \exp\left(\mathrm{j}\, 2\pi q \frac{t}{T}\right), \quad t \in \mathbb{R}, \tag{2.1}$$

yields the signal model

$$\boxed{x_b(t) = \sum_{q=0}^{N-1} \underbrace{X_{b,q}}_{\text{Data}} \; \psi_{b,q}(t), \quad t \in \mathbb{T}_b} \tag{2.2}$$

and

$$\boxed{x(t) = \sum_{b=-\infty}^{\infty} p_{\mathbb{T}_b}(t) \underbrace{\sum_{q=0}^{N-1} X_{b,q} \exp\left(\mathrm{j}\, 2\pi q \frac{t}{T}\right),}_{b\text{-th channel use}} \quad t \in \mathbb{R},} \tag{2.3}$$

with

$$p_{\mathbb{T}_b}(t) = \begin{cases} 1 \; ; & t \in \mathbb{T}_b \\ 0 \; ; & \text{otherwise.} \end{cases} \tag{2.4}$$

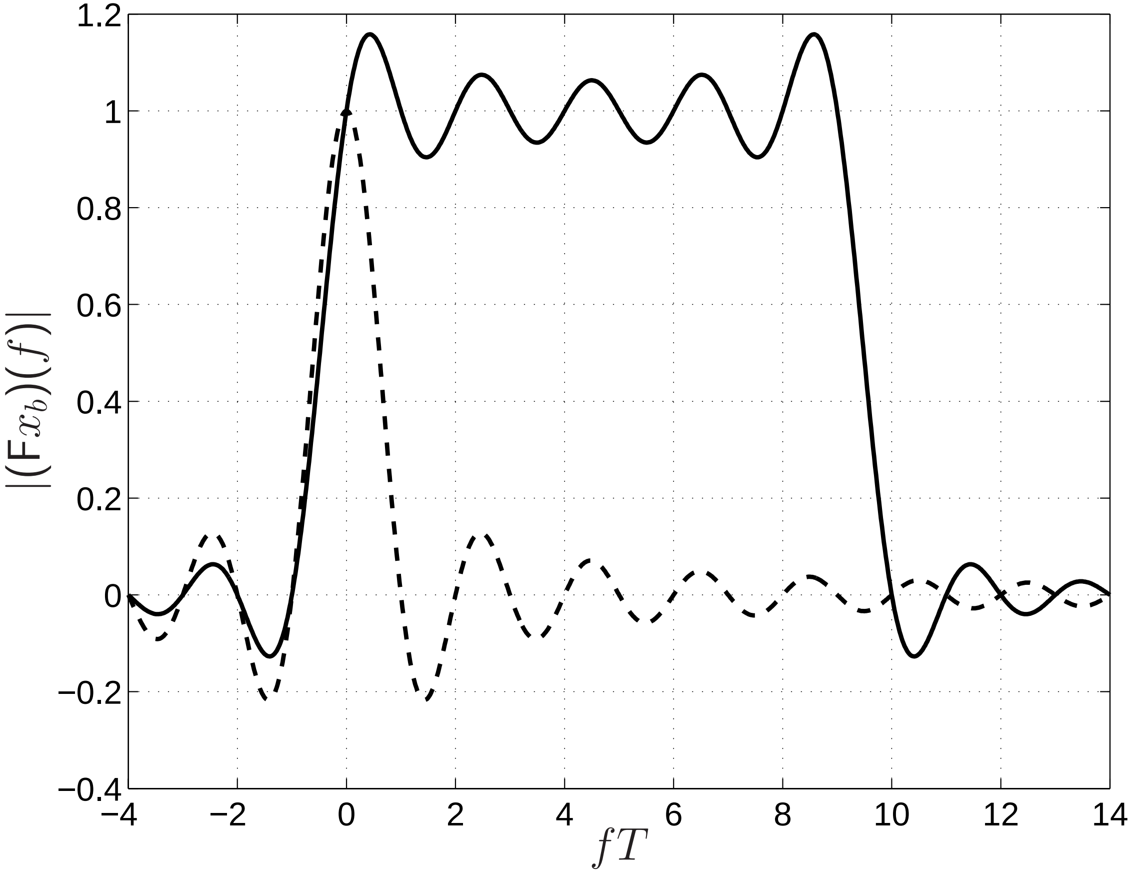

Fig. 2.1. Spectrum of a single carrier signal and the full OFDM multicarrier signal $x_b(t)$ with $N = 10$ and $X_{b,0} = \cdots = X_{b,N-1}$

2.1 OFDM Demodulation (AWGN)

Since the carrier functions are orthogonal, the signal demodulation at the receiver can be based on the orthogonal projection $\mathsf{P}_{b,q} \triangleq \psi_{b,q}(t)\langle \bullet, \psi_{b,q}\rangle_{\mathbb{T}_b}$ of the receiver signal[1]

$$y(t) = \sum_{b=-\infty}^{\infty} \sum_{q=0}^{N-1} X_{b,q}\psi_{b,q}(t) + \eta(t). \tag{2.5}$$

The required definition of the inner product can be obtained as

$$\langle g, h\rangle_{\mathbb{T}_b} \triangleq \frac{1}{T}\int_{bT}^{(b+1)T} g(t)\, h^*(t)\, \mathrm{d}t. \tag{2.6}$$

The projected signal equals $\hat{x}_{b,q}(t) = (\mathsf{P}_{b,q}y)(t)$ and the estimate of the transmitted data $X_{b,q}$ reads[2]

$$\boxed{\hat{X}_{b,q} = \langle y, \psi_{b,q}\rangle_{\mathbb{T}_b} = X_{b,q} + \langle \eta, \psi_{b,q}\rangle_{\mathbb{T}_b}.} \tag{2.7}$$

[1] The assumption of an AWGN transmission channel leads to $\varphi_{b,q}(t) = \psi_{b,q}(t)$.

[2] $\langle y, \psi_{b,q}\rangle_{\mathbb{T}_b} = \sum_{b'=-\infty}^{\infty}\sum_{q'=1}^{N} X_{b',q'}\langle \psi_{b',q'}, \psi^*_{b,q}\rangle_{\mathbb{T}_b} + \langle \eta, \psi_{b,q}\rangle_{\mathbb{T}_b} = \sum_{b'=-\infty}^{\infty}\sum_{q'=1}^{N} X_{b',q'}\delta_{b',b}\delta_{q',q} + \eta_{b,q} = X_{b,q} + \eta_{b,q}$.

2.2 OFDM Demodulation (Dispersive Channels)

Given dispersive transmission channels[3]

$$h(t) = \sum_{\ell=0}^{L-1} h_\ell \, \delta\left(t - \ell \frac{T}{N}\right), \tag{2.8}$$

the delayed reception of signals $h_\ell x_{b,q} \, \psi_q(t - \ell \frac{T}{N} - bT)$ leads to interference effects between adjacent channel uses $\ldots, \mathbb{T}_{b-1}, \mathbb{T}_b, \mathbb{T}_{b+1}, \ldots$ (Fig. 2.2).

A simple interference mitigation is based on a cyclic extension of the signal model at the transmitter. To this end, $T_\mathsf{s} = T + T_\mathsf{p}$ and $\psi_{b,q}(t), \varphi_{b,q}(t) \in \mathbb{T}_b = [-T_\mathsf{p} + bT_\mathsf{s}, bT_\mathsf{s} + T[$ are defined as (e. g. $b = 0$)

$$\psi_{0,q}(t) \;=\; p_{\mathbb{T}_0}(t) \exp\left(\mathrm{j}\, 2\pi q \frac{t}{T}\right) \tag{2.9}$$

$$\varphi_{0,q}(t) \;=\; \begin{cases} 0 & ; \; -T_\mathsf{p} \le t < 0 \\ \exp\left(\mathrm{j}\, 2\pi q \frac{t}{T}\right) & ; \; 0 \le t < T \\ 0 & ; \; t \notin \mathbb{T}_b. \end{cases} \tag{2.10}$$

[3] The equidistant power delay profile is assumed to be aligned with the time instants $\ell \frac{T}{N} = \ell \frac{1}{B}$.

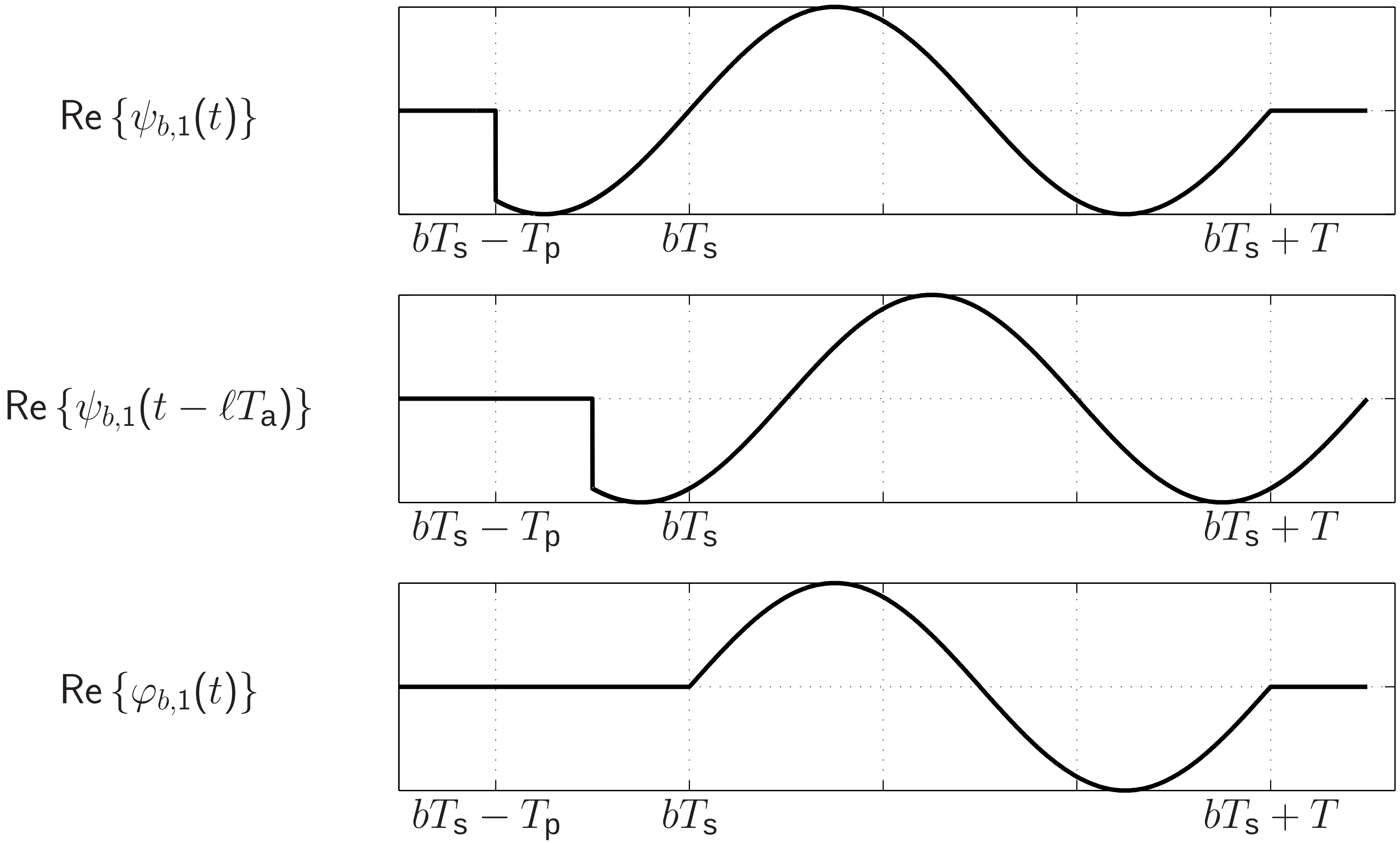

Fig. 2.2. Cyclic Extension of an OFDM signal ($\ell T_\mathsf{a} < T_\mathsf{p}$)

If the guard intervall is larger or equal to the maximum channel delay time, $T_{\mathsf{p}} \geq (L-1)\frac{T}{N}$, the signal demodulation at the receiver is again given by the orthogonal projection $\mathsf{P}_{b,q} \triangleq \psi_{b,q}(t)\langle \bullet, \psi_{b,q}\rangle_{\mathbb{T}_b}$ of the received signal

$$\boxed{y(t) = \sum_{b=-\infty}^{\infty} \sum_{q=0}^{N-1} \sum_{\ell=0}^{L-1} h_\ell X_{b,q} \psi_{b,q}\left(t - \ell\frac{T}{N}\right) + \eta(t).} \tag{2.11}$$

The required definition of the inner product can be obtained as

$$\langle g, h\rangle_{\mathbb{T}_b} \triangleq \frac{1}{T} \int_{-T_{\mathsf{p}}+bT_{\mathsf{s}}}^{bT_{\mathsf{s}}+T} g(t)\, h^*(t)\, \mathsf{d}t. \tag{2.12}$$

The projected signal equals $\hat{x}_{b,q}(t) = (\mathsf{P}_{b,q}y)(t)$ and the estimate of the transmitted data $X_{b,q}$ reads

$$\boxed{\hat{X}_{b,q} = \langle y, \varphi_{b,q}(t)\rangle_{\mathbb{T}_b} = N \cdot X_{b,q}\, H\left(\frac{q}{T}\right) + \langle \eta, \varphi_{b,q}\rangle_{\mathbb{T}_b},} \tag{2.13}$$

with $H\left(\frac{q}{T}\right) = \frac{1}{N}\left(\mathsf{F}h\right)(f)\big|_{f=q\frac{B}{N}} = \frac{1}{N}\sum_{\ell=0}^{L-1} h_\ell \exp\left(\mathsf{j}\, 2\pi \frac{q\ell}{N}\right)$ as the equidistant sampled channel transmission function $\frac{1}{N}(\mathsf{F}h)(f)$ in frequency domain.

Proof

Since $\psi_{b,q}(t-\Delta t) = \exp\left(-\mathsf{j}\,2\pi q \frac{\Delta t}{T}\right)\psi_{b,q}(t)$ if $bT_\mathsf{s} \leq t < bT_\mathsf{s} + T$ and $0 \leq \Delta t \leq T_\mathsf{p}$, and $\psi_{b,q}(t) = \varphi_{b,q}(t)$ if $bT_\mathsf{s} \leq t < bT_\mathsf{s} + T$,

$$\begin{aligned}
\hat{X}_{b,q} &= \sum_{q'=0}^{N-1}\sum_{\ell=0}^{L-1} X_{b,q'} h_\ell \int_{bT_\mathsf{s}}^{bT_\mathsf{s}+T} \psi_{b,q'}\left(t-\ell\frac{T}{N}\right)\varphi^*_{b,q}(t)\,\mathsf{d}t + \eta_{b,q} \\
&= \sum_{q'=0}^{N-1}\sum_{\ell=0}^{L-1} X_{b,q'} h_\ell\, \mathsf{e}^{-\mathsf{j}\,2\pi q \frac{\ell T}{TN}} \left\langle \psi_{b,q'}, \psi_{b,q}\right\rangle_{\mathbb{T}_b} + \eta_{b,q} \\
&= X_{b,q}\sum_{\ell=0}^{L-1} h_\ell\, \mathsf{e}^{-\mathsf{j}\,2\pi \frac{q\ell}{N}} + \eta_{b,q} = N \cdot X_{b,q} H\left(\frac{q}{T}\right) + \eta_{b,q}, \quad \text{with}
\end{aligned}$$

$$H\left(\frac{q}{T}\right) = \frac{1}{N}\sum_{\ell=0}^{L-1} h_\ell\, \mathsf{e}^{-\mathsf{j}\,2\pi \frac{q\ell}{N}} = \frac{1}{N}\boldsymbol{e}_{q+1}^\mathsf{T}\boldsymbol{F}\boldsymbol{h} \quad \text{(DFT)} \tag{2.14}$$

and $\boldsymbol{e}_{q+1}^\mathsf{T}\boldsymbol{F}$ equal to the $(q+1)$-th row of the Fouriermatrix $\boldsymbol{F} \in \mathbb{C}^{N\times N}$ und $\boldsymbol{h} = [h_0, h_1, \ldots, h_{L-1}, 0, \ldots, 0]^\mathsf{T} \in \mathbb{C}^N$. □

2.3 Realization of the Demodulation Function

The demodulation function $\langle y, \varphi_{b,q} \rangle_{\mathbb{T}_b}$ can be realized either by means of a correlator or a matched filter. For the derivation of the matched filter solution the inner product is rewritten as follows:

$$\begin{aligned}
\langle y, \varphi_{b,q} \rangle_{\mathbb{T}_b} &= \int_{-T_\mathsf{p}+bT_\mathsf{s}}^{bT_\mathsf{s}+T} y(t)\varphi^*_{b,q}(t)\,\mathsf{d}t \\
&= \int_{-T_\mathsf{p}+bT_\mathsf{s}}^{bT_\mathsf{s}+T} y(t)\psi^*_{b,q}(t) g_{\mathbb{T}_b}(t)\,\mathsf{d}t = \int_{-\infty}^{\infty} y(t)\,\mathsf{exp}\left(-\,\mathsf{j}\,2\pi q \frac{t}{T}\right) g_{\mathbb{T}_b}(t)\,\mathsf{d}t,
\end{aligned}$$

with

$$g_{\mathbb{T}_b}(t) = \begin{cases} 0 \;;\; -T_\mathsf{p} \leq t < 0 \\ 1 \;;\; \quad 0 \leq t < T \\ 0 \;;\; \qquad \text{sonst}, \end{cases} \tag{2.15}$$

i. e.

$$\boxed{\langle y, \varphi_{b,q} \rangle_{\mathbb{T}_b} = \left[y(t) \cdot \mathsf{exp}\left(-\,\mathsf{j}\,2\pi q \tfrac{t}{T}\right)\right] * g^*_{\mathbb{T}_0}(-t)\Big|_{t=bT_\mathsf{s}}.} \tag{2.16}$$

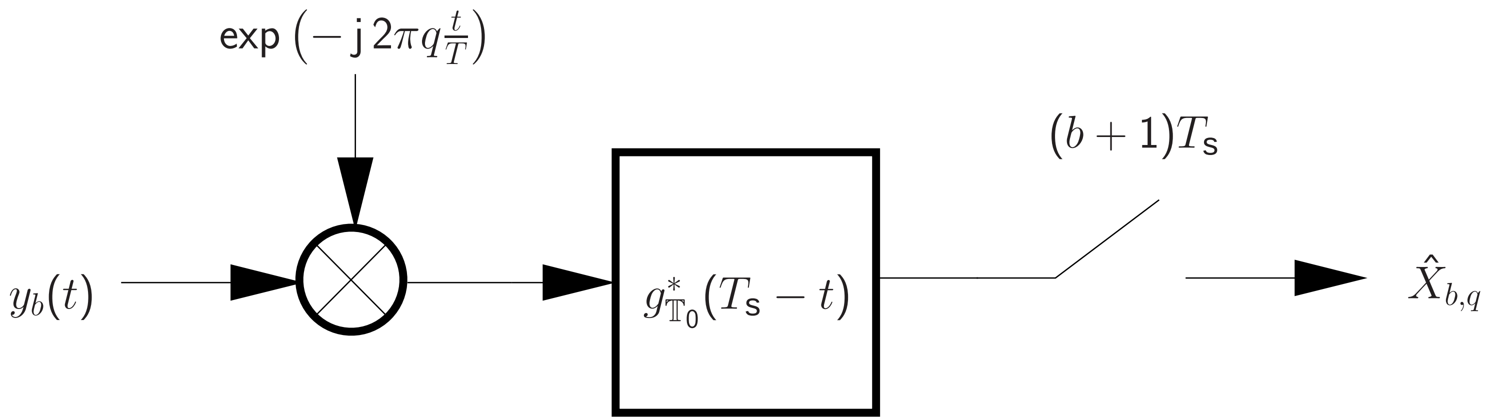

Fig. 2.3. Realization of the Demodulation Function by Correlation / Matched Filter

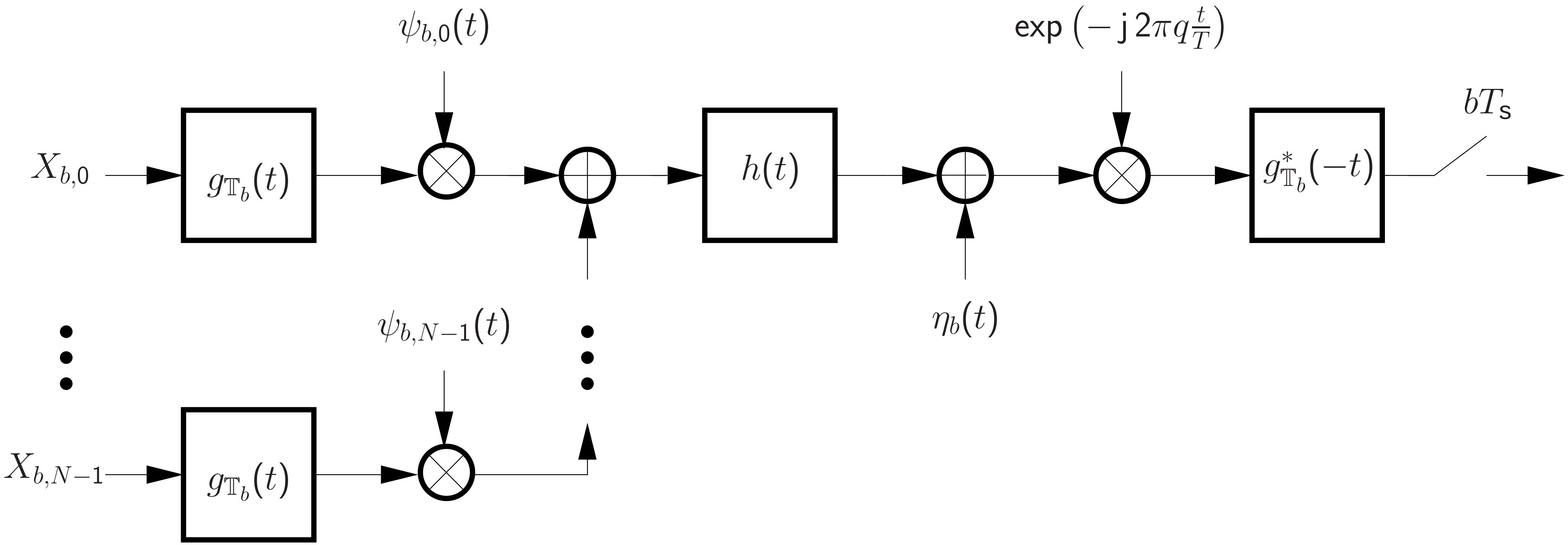

Fig. 2.4. Principle of Modulation / Demodulation in OFDM

2.4 OFDM using the Discrete Fourier Transform (DFT/IDFT)

The sampling of the modulated signal $x_b(t)$ with sampling rate $B = N/T$, i. e. $T_\mathsf{a} = T/N$, yields the sequence of discrete-time signals

$$x_{b,n} = x_b\left(bT_\mathsf{s} + n\frac{T}{N}\right) = \sum_{q=0}^{N-1} \gamma_b^q \, X_{b,q} \exp\left(\mathrm{j}\,2\pi \frac{qn}{N}\right), \tag{2.17}$$

with

$$n \in \{-N_\mathsf{p}, \dots, N-1\}, \quad \gamma_b \triangleq \exp\left(\mathrm{j}\,2\pi b \frac{T_\mathsf{p}}{T}\right) \quad \text{und} \quad N_\mathsf{p} \triangleq \frac{T_\mathsf{p}}{T_\mathsf{a}}. \tag{2.18}$$

The sequence $\{x_{b,n}\}_{n=-N_\mathsf{p}}^{N-1}$ obviously results from the inverse DFT of $\boldsymbol{X}_b \triangleq [X_{b,0}, \dots, X_{b,N-1}]^\mathsf{T}$ in $\boldsymbol{x}_b = [x_{b,0}, \dots, x_{b,-1}]^\mathsf{T}$, and a subsequent cyclic extension of $\boldsymbol{x}_b$ to the elements $\{x_{b,n}\}_{q=-N_\mathsf{p}}^{-1}$, i. e.

$$\boxed{\boldsymbol{x}_b = \boldsymbol{F}^\mathsf{H} \boldsymbol{X}_b \quad \text{and} \quad x_{b,n} = x_{b,\mathrm{mod}_N(n)} = x_{b,N-n}, \quad n = -N_\mathsf{p}, \dots, -1.} \tag{2.19}$$

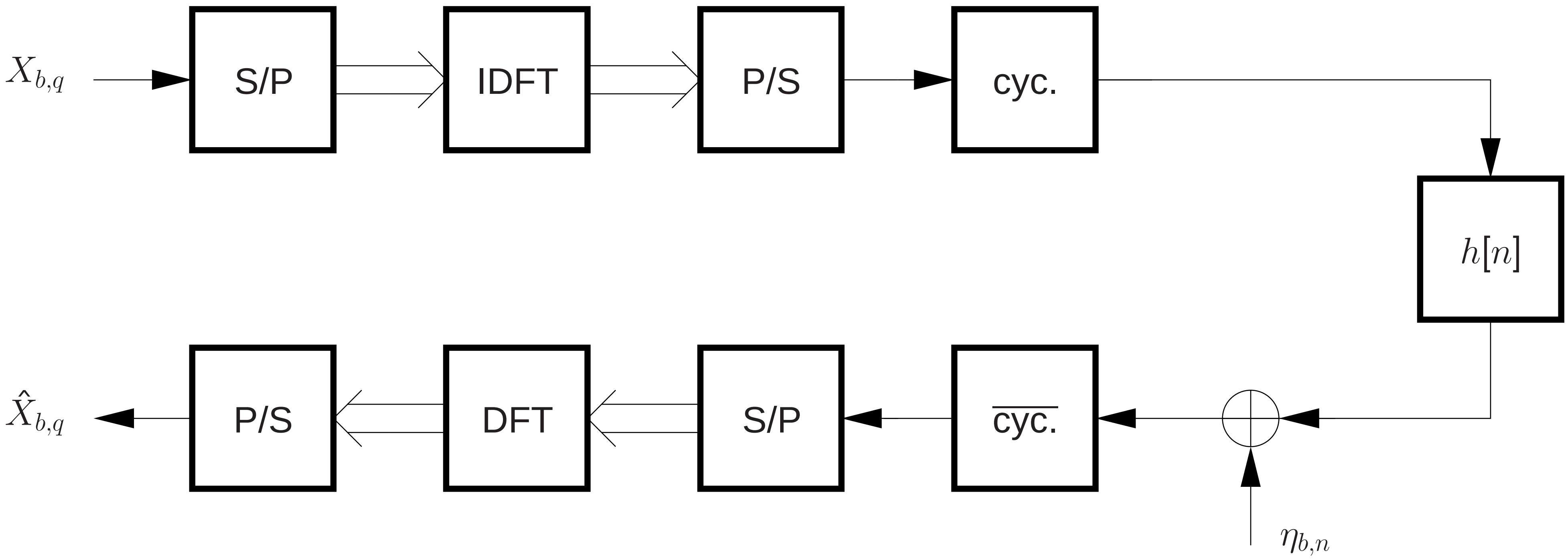

Fig. 2.5. OFDM using IDFT und DFT

2.5 Dispersive Transmission and Cyclic Convolution

Given a dispersive transmission channel $h[n] = \sum_{\ell=1}^{L-1} h_\ell \, \delta[k-\ell]$, and the cyclic extension of the transmitted signals $x_b[n]$, the received signals are given by

$$y_{b,n} = \sum_{\ell=1}^{L-1} h_\ell \, \delta_{n-\ell} * \sum_{k=-N_\mathsf{p}}^{N-1} x_{b,k} \, \delta_{n-k} = \sum_{\ell=1}^{L-1} \sum_{k=-N_\mathsf{p}}^{N} h_\ell \, x_{b,k} \, \delta_{n-\ell-k}, \quad (2.20)$$

subject to $L - 1 \leq N_\mathsf{p}$, i. e.

$$y_{b,n} = \sum_{\ell=1}^{L-1} h_\ell \, x_{b,n-\ell} = \sum_{\ell=1}^{L-1} h_\ell \, x_{b,\mathrm{mod}_N(n-\ell)}, \quad n-\ell \in \{-N_\mathsf{p}, \ldots, N-1\}. \quad (2.21)$$

Given the discrete-time cyclic convolution, the following Theorem is inherited from the continuous-time case:

$$\boxed{y_{b,n} = h_n \circ x_{b,n} \quad \Leftrightarrow \quad \tfrac{1}{N} Y_{b,q} = H_q \cdot X_{b,q}, \qquad n,q = 0,\ldots,N-1} \quad (2.22)$$

and H_q with respect to (2.14).

Example

Given the channel impulse response function $h[n] = \sum_{\ell=0}^{2} h_\ell \, \delta[n-\ell]$, $N = 4$ and $N_\mathsf{p} = 2$, the received vector of signal equals

$$
\begin{bmatrix} y_{b,-2} \\ y_{b,-1} \\ y_{b,0} \\ y_{b,1} \\ y_{b,2} \\ y_{b,3} \end{bmatrix}
=
\begin{bmatrix} h_0 & & & & & \\ h_1 & h_0 & & & & \\ h_2 & h_1 & h_0 & & & \\ & h_2 & h_1 & h_0 & & \\ & & h_2 & h_1 & h_0 & \\ & & & h_2 & h_1 & h_0 \end{bmatrix}
\begin{bmatrix} x_{b,-2} \\ x_{b,-1} \\ x_{b,0} \\ x_{b,1} \\ x_{b,2} \\ x_{b,3} \end{bmatrix}
+
\begin{bmatrix} h_2 & h_1 \\ & h_1 \\ & \\ & \\ & \\ & \end{bmatrix}
\begin{bmatrix} x_{b-1,2} \\ x_{b-1,3} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
+
\begin{bmatrix} \eta_{b,-2} \\ \eta_{b,-1} \\ \eta_{b,0} \\ \eta_{b,1} \\ \eta_{b,2} \\ \eta_{b,3} \end{bmatrix}
$$

After cyclic extension $x_{b,-n} = x_{b,N-n}$, $n \leq N_\mathsf{p}$, viz. $x_{b,-2} = x_{b,2}$ and $x_{b,-1} = x_{b,3}$, the deletion of $y_{b,-2}$ and $y_{b,-1}$, and some reordering,

$$
\begin{bmatrix} y_{b,0} \\ y_{b,1} \\ y_{b,2} \\ y_{b,3} \end{bmatrix}
=
\begin{bmatrix} h_0 & & h_2 & h_1 \\ h_1 & h_0 & & h_2 \\ h_2 & h_1 & h_0 & \\ & h_2 & h_1 & h_0 \end{bmatrix}
\begin{bmatrix} x_{b,0} \\ x_{b,1} \\ x_{b,2} \\ x_{b,3} \end{bmatrix}
+
\begin{bmatrix} \eta_{b,0} \\ \eta_{b,1} \\ \eta_{b,2} \\ \eta_{b,3} \end{bmatrix}
\tag{2.23}
$$

$$\boxed{\breve{\boldsymbol{y}}_b = \boldsymbol{H}\breve{\boldsymbol{x}}_b + \breve{\boldsymbol{z}}_b + \breve{\boldsymbol{\eta}}_b \quad \Leftrightarrow \quad \boldsymbol{y}_b = \breve{\boldsymbol{H}}\boldsymbol{x}_b + \boldsymbol{\eta}_b.} \tag{2.24}$$

The cyclic extension $\breve{\boldsymbol{x}}_b$ ($\rightarrow$ cyclic convolution) causes the convolution matrix $\boldsymbol{H} \in \mathbb{C}^{6\times 6}$ to appear as the cyclic convolution matrix $\breve{\boldsymbol{H}} \in \mathbb{C}^{4\times 4}$:

$$[\breve{\boldsymbol{H}}]_{n,k} = h_{\mathrm{mod}_N(n-k)}. \tag{2.25}$$

The DFT of (2.23) yields

$$\boldsymbol{y}_b = \breve{\boldsymbol{H}}\boldsymbol{x}_b + \boldsymbol{\eta}_b \quad \Leftrightarrow \quad \boldsymbol{Y}_b = \frac{1}{N}\boldsymbol{F}\breve{\boldsymbol{H}}\boldsymbol{F}^{\mathsf{H}}\boldsymbol{X}_b + \frac{1}{N}\boldsymbol{F}\boldsymbol{\eta}_b. \tag{2.26}$$

Finally, the cyclic structur of $\breve{\boldsymbol{H}}$ leads to

$$\boxed{\boldsymbol{F}\breve{\boldsymbol{H}}\boldsymbol{F}^{\mathsf{H}} = N^2 \cdot \mathsf{diag}\left[H_0, H_1, \ldots, H_{N-1}\right],} \tag{2.27}$$

i. e. die matrix of eigenvectors of a cyclic Toeplitzmatrix ist unitary and a priori known as the transformation matrix of the IDFT.[4]

[4]The eigenvalues equal to the spectrum of the transformed (DFT) sequence of $h_0, \ldots, h_{N-1}$.

Proof

Given $\breve{\boldsymbol{h}}_k = \breve{\boldsymbol{H}}\boldsymbol{e}_{k+1}$, the sequence of elements from $\breve{\boldsymbol{h}}_k$ can be interpreted as the cyclic convolution of $\boldsymbol{h} = [h_0, \ldots, h_{L-1}, 0, \ldots, 0]^\mathsf{T}$ with the shift operator $\delta_{k,n}$, i. e. $[\breve{\boldsymbol{h}}_k]_n = \sum_{\ell=1}^{L-1} h_\ell \, \delta_{k,\mathrm{mod}(n-\ell)} = h_{\mathrm{mod}(n-k)}$. Applying the convolution Theorem,

$$\underbrace{h_{\mathrm{mod}(n-k)}}_{\text{Time domain}} \Leftrightarrow \underbrace{H_q \exp\left(-\mathrm{j}\,2\pi\frac{qk}{N}\right)}_{\text{Frequency domain}}, \tag{2.28}$$

leads to

$$\begin{aligned}\frac{1}{N}\boldsymbol{F}\breve{\boldsymbol{h}}_k &= \mathrm{diag}\left[\mathrm{e}^{-\mathrm{j}\,2\pi\frac{qk}{N}}\right]_{q=0}^{N-1}[H_0, \ldots, H_{N-1}]^\mathsf{T} \\ &= \mathrm{diag}\,[H_0, \ldots, H_{N-1}]\left[1, \ldots, \mathrm{e}^{-\mathrm{j}\,2\pi\frac{(N-1)k}{N}}\right]^\mathsf{T} = \mathrm{diag}\,[H_0, \ldots, H_{N-1}]\,\boldsymbol{F}\boldsymbol{e}_{k+1}.\end{aligned}$$

and thus $\frac{1}{N}\boldsymbol{F}\breve{\boldsymbol{H}} = \frac{1}{N}\boldsymbol{F}[\breve{\boldsymbol{h}}_0, \ldots, \breve{\boldsymbol{h}}_{N-1}] = \mathrm{diag}\,[H_0, \ldots, H_{N-1}]\,\boldsymbol{F}$.

Finally, multiplication by $\boldsymbol{F}^{-1} = \frac{1}{N}\boldsymbol{F}^\mathsf{H}$ yields

$$\boldsymbol{F}\breve{\boldsymbol{H}}\boldsymbol{F}^\mathsf{H} = N^2 \cdot \mathrm{diag}\,[H_0, \ldots, H_{N-1}]\,.$$

$\square$

Principles of Channel Estimation for OFDM

3. Channel Modelling for OFDM Systems

A simplified model of the channel transmission function (operator kernel) generally reads

$$\boxed{H(f,t) = \sum_{\ell=1}^{L_\infty} \alpha_\ell \, \mathsf{e}^{-\mathsf{j}\,2\pi f \tau_\ell} \, \mathsf{e}^{\mathsf{j}\,2\pi \nu_\ell t},} \tag{3.1}$$

where $f \in \mathbb{R}$ and $t \in \mathbb{R}$ represent the frequency and time domain, and $\alpha_\ell \in \mathbb{C}$, $\tau_\ell \in \mathbb{R}$, and $\nu_\ell \in \mathbb{R}$ denotes the ℓ-th attenuation coefficient, time delay, and Doppler frequency, respectively.

In general the following mixed representation between channel transmission function and impulse response function is used:

$$h(\tau,t) = (\mathsf{F}^{-1}\mathsf{H})(\tau,t), \tag{3.2}$$

i. e. the inverse Fourier transform of $H(f,t)$ with respect to f- domain:

$$\boxed{h(\tau,t) = \sum_{\ell=1}^{L_\infty} \underbrace{\left(\alpha_\ell \, \mathsf{e}^{\mathsf{j}\,2\pi \nu_\ell t}\right)}_{h_\ell(t)} \delta(\tau - \tau_\ell).} \tag{3.3}$$

3.1 Stochastic Modelling

The received signal $y(t)$ can be obtained as (infinite) multiple copies of the transmitted signal $x(t)$:

$$y(t) = \int_{-\infty}^{\infty} H(f,t)\, \underbrace{(\mathsf{F}x)(f)}_{X(f)}\, \mathsf{exp}(\mathsf{j}\, 2\pi f t)\, \mathsf{d}f \tag{3.4}$$

$$= \sum_{\ell=1}^{L_\infty} \underbrace{\left(\alpha_\ell\, \mathsf{e}^{\mathsf{j}\, 2\pi\nu_\ell t}\right)}_{h_\ell(t)}\, x(t-\tau_\ell) \tag{3.5}$$

Due to the limitted bandwidth, the originally large number of time delays τ_ℓ is replaced by a finite number L of resolvable time delays $\hat{\tau}_\ell$:[1]

$$\boxed{y(t) = \sum_{\ell=1}^{L} h_\ell(t)\, x(t-\hat{\tau}_\ell),} \tag{3.6}$$

where $h_\ell(t) = \sum\limits_{l\in\{\tau_l\approx\hat{\tau}_\ell\}} \alpha_l\, \mathsf{e}^{\mathsf{j}\, 2\pi\nu_l t}$ is generally time variant.

[1]The resolution of time delays depends on the available bandwidth of the communication system. In the sequel, τ_ℓ instead of $\hat{\tau}_\ell$ is used.

Since generally the time variant function is unknown, a stochastic model is applied, where $h_\ell(t)$ represents an realization of the random process $h_{\ell,t}$. Instead of the properties of function $h_\ell : \mathbb{R} \to \mathbb{C}, t \mapsto h_\ell(t)$, the statistical properties of the random process $h_{\ell,t} : (\Omega, \Sigma, P) \to \mathbb{C} : \omega \mapsto h_\ell(t)$ is discussed.

Depending on the multipath scenario of the given communication system different probability distributions and autocorrelation function of $h_{\ell,t}$ are assumed. We constrain ourselfs on Rayleigh distribution and Clarke-Jakes autocorrelation function:

$$\boxed{p_{|h_{\ell,t}|}(\xi) = \frac{\xi}{2\sigma_{h,\ell}^2} \exp\left(-\frac{\xi^2}{2\sigma_{h,\ell}^2}\right),} \tag{3.7}$$

where

$$\boxed{\sigma_{h,\ell}^2 = \mathsf{E}\left[h_{\ell,t} h_{\ell,t}^*\right],} \tag{3.8}$$

and[2]

$$\boxed{\mathsf{E}\left[h_{\ell,t} h_{\ell,t'}^*\right] = \sigma_{h,\ell}^2 J_0\left(2\pi f_{\mathsf{D}}(t-t')\right).} \tag{3.9}$$

[2] $J_0(\xi) = \frac{1}{2\pi}\int_0^{2\pi} \cos(\xi \sin\alpha)\, \mathrm{d}\,\alpha$.

Wide Sense Stationary Uncorrelated Scattering (WSSUS)

Given the correspondence between the transfer function and the spreading function and assumption of discrete kernels,

$$\boxed{H(f,t) = \sum_{\ell=1}^{L_\infty} \alpha_\ell \, \mathrm{e}^{-\mathrm{j}\,2\pi f \tau_\ell} \, \mathrm{e}^{\mathrm{j}\,2\pi \nu_\ell t} \quad \Leftrightarrow \quad h(\tau,\nu) = \sum_{\ell=1}^{L_\infty} \alpha_\ell \, \delta(\tau-\tau_\ell)\, \delta(\nu-\nu_\ell),}$$

we obtain stationarity in the wide sense,

$$\boxed{\mathrm{E}\left[H(f,t)H^*(f',t')\right] = R(f-f',t-t'),} \tag{3.10}$$

and uncorrelated scattering,

$$\boxed{\mathrm{E}\left[h(\tau,\nu)h^*(\tau',\nu')\right] = r(\tau-\tau',\nu-\nu'),} \tag{3.11}$$

where

$$r(\tau-\tau',\nu-\nu') = S(\tau,\nu)\,\delta(\tau-\tau')\,\delta(\nu-\nu'), \tag{3.12}$$

with the scattering function

$$\boxed{S(\tau,\nu) = \sum_{\ell=1}^{L_\infty} \mathrm{E}\left[|\alpha_\ell|\right]^2 \; \delta(\tau-\tau_\ell)\,\delta(\nu-\nu_\ell).} \tag{3.13}$$

4. Statistical Theory of Reference Data Assisted Estimation

An optimal estimator $\boldsymbol{G} : \boldsymbol{y} \mapsto \hat{\boldsymbol{h}}$ of unknown parameters $\boldsymbol{h} \in \mathbb{C}^N$ based on given reference data $x_1, \ldots, x_M \in \mathbb{C}$ and observations $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_M \in \mathbb{C}^N$, depending on $\boldsymbol{h}$ and the reference data, can be based on the minimization of the Risk Functional

$$\boxed{\min \mathsf{E}\left[L(\boldsymbol{y}, \boldsymbol{h}; x_1, \ldots, x_M)\right] \approx \min \sum_{i=1}^{M} \frac{1}{M} L(\boldsymbol{y}_i, \boldsymbol{h}_i; x_1, \ldots, x_M),} \tag{4.1}$$

where $L(\boldsymbol{y}, \boldsymbol{h}; x_1, \ldots, x_M)$ represents the Loss Function which determines the nature of the estimation approach:

- Maximum Likelihood (ML) estimation[1] by $\min_{\boldsymbol{h}} \mathsf{E}\left[\bullet\right]$ of

$$L(\boldsymbol{y}_i; \boldsymbol{h}, x_i) = -\log f_{\boldsymbol{y}|x}(\boldsymbol{y}_i \mid x_i; \boldsymbol{h}), \tag{4.2}$$

- Minimum Mean Square Error estimation[2] (MSE) by $\min_{\boldsymbol{G}} \mathsf{E}\left[\bullet\right]$ of

$$L(\boldsymbol{y}, \boldsymbol{h}; x_1, \ldots, x_M) = \left(\boldsymbol{h} - \boldsymbol{G}(\boldsymbol{y}; x_1, \ldots, x_M)\right)^2. \tag{4.3}$$

[1] ML estimation is based on the empirical cost function. The unknown $\boldsymbol{h}$ is modelled as a **deterministic vector** which describes a chosen probability distribution.

[2] Minimum MSE estimation is based on the bayesian cost function. The unknown $\boldsymbol{h}$ is modelled as a **random vector** which describes the a chosen mapping between the reference data and the observations.

4.1 ML Estimation ($\rightarrow$ LS)

Applying the signal model $\boldsymbol{y} = \boldsymbol{f}(x; \boldsymbol{h}) + \boldsymbol{\eta}$ and the gaussian assumption

$$\boldsymbol{y} - \boldsymbol{f}(x; \boldsymbol{h}) \sim \mathsf{N}\left(\boldsymbol{0}, \boldsymbol{C}_{\eta,\eta}\right), \qquad \boldsymbol{C}_{\eta,\eta} = \mathsf{E}\left[\boldsymbol{\eta}\boldsymbol{\eta}^{\mathsf{H}}\right], \tag{4.4}$$

the minimization of the empirical cost function leads to[3]

$$\boxed{\hat{\boldsymbol{h}}_{\mathsf{ML}} = \mathsf{argmax}_{\boldsymbol{h}} \sum_{i=1}^{M} \log p_{\boldsymbol{y}|\mathsf{x}}(\boldsymbol{y}_i \mid x_i; \boldsymbol{h}) = \mathsf{argmin}_{\boldsymbol{h}} \sum_{i=1}^{M} \|\boldsymbol{y}_i - \boldsymbol{f}(x_i; \boldsymbol{h})\|^2_{\boldsymbol{C}^{-1}_{\eta,\eta}}.} \tag{4.5}$$

Given a linear model $\boldsymbol{y} = \boldsymbol{h}x + \boldsymbol{\eta}$, observations $\bar{\boldsymbol{y}} = [\boldsymbol{y}_1^{\mathsf{T}}, \ldots, \boldsymbol{y}_M^{\mathsf{T}}]^{\mathsf{T}}$, reference data $\bar{\boldsymbol{x}} = [x_1, \ldots, x_M]^{\mathsf{T}}$, and noise vector $\bar{\boldsymbol{\eta}} = [\boldsymbol{\eta}_1^{\mathsf{T}}, \ldots, \boldsymbol{\eta}_M^{\mathsf{T}}]^{\mathsf{T}}$, i. e.

$$\bar{\boldsymbol{y}} = \underbrace{(\bar{\boldsymbol{x}} \otimes \boldsymbol{I}_N)}_{\bar{\boldsymbol{X}}} \boldsymbol{h} + \bar{\boldsymbol{\eta}} \tag{4.6}$$

the solution of (4.5) can be obtained as

$$\boxed{\hat{\boldsymbol{h}}_{\mathsf{ML}} = \left(\bar{\boldsymbol{X}}^{\mathsf{H}} \boldsymbol{C}_{\bar{\eta},\bar{\eta}}{}^{-1} \bar{\boldsymbol{X}}\right)^{-1} \bar{\boldsymbol{X}}^{\mathsf{H}} \boldsymbol{C}_{\bar{\eta},\bar{\eta}}{}^{-1} \bar{\boldsymbol{y}}.} \tag{4.7}$$

[3] If the elements of $\boldsymbol{\eta}$ are i.i.d. the ML estimation simplifies to the Least Squares (LS) estimation.

Remark

$$\bar{\boldsymbol{y}} = (\bar{\boldsymbol{x}} \otimes \boldsymbol{I}_N)\,\boldsymbol{h} + \bar{\boldsymbol{\eta}}$$

$$\Downarrow$$

$$\begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_M \end{bmatrix} = \left[\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{bmatrix} \otimes \boldsymbol{I}_N\right] \boldsymbol{h} + \begin{bmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \\ \vdots \\ \boldsymbol{\eta}_M \end{bmatrix}$$

$$= \begin{bmatrix} x_1 \cdot \boldsymbol{I}_N \\ x_2 \cdot \boldsymbol{I}_N \\ \vdots \\ x_M \cdot \boldsymbol{I}_N \end{bmatrix} \boldsymbol{h} + \begin{bmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \\ \vdots \\ \boldsymbol{\eta}_M \end{bmatrix},$$

where $\otimes$ denotes the Kronecker product of two matrices.[4]

[4]Given $\boldsymbol{A} \in \mathbb{C}^{M_a \times N_a}$ and $\boldsymbol{B} \in \mathbb{C}^{M_b \times N_b}$,

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{bmatrix} a_{1,1}\boldsymbol{B} & a_{1,2}\boldsymbol{B} & \cdots & a_{1,N_a}\boldsymbol{B} \\ a_{2,1}\boldsymbol{B} & a_{2,2}\boldsymbol{B} & \cdots & a_{2,N_a}\boldsymbol{B} \\ \vdots & \vdots & & \vdots \\ a_{M_a,1}\boldsymbol{B} & a_{M_a,2}\boldsymbol{B} & \cdots & a_{M_a,N_a}\boldsymbol{B} \end{bmatrix} \in \mathbb{C}^{M_a M_b \times N_a N_b}.$$

4.2 MMSE Estimation (Wiener Filter)

Applying the above signal model the minimization of the bayesian cost function leads to the Conditional Mean Estimator (CME)

$$\hat{\boldsymbol{h}}_{\mathsf{WF}} = \boldsymbol{G}(\bar{\boldsymbol{y}}; \bar{\boldsymbol{x}}) = \mathsf{E}\left[\boldsymbol{h} \mid \bar{\mathsf{y}} = \bar{\boldsymbol{y}}; \bar{\boldsymbol{x}}\right]. \tag{4.8}$$

In the case of a linear estimation $\hat{\boldsymbol{h}}_{\mathsf{WF}} = \boldsymbol{G}(\bar{\boldsymbol{y}}; \bar{\boldsymbol{x}}) = \boldsymbol{G}\bar{\boldsymbol{y}}$, the solution of (4.8) can be obtained as[5]

$$\begin{aligned}
\hat{\boldsymbol{h}}_{\mathsf{WF}} &= \mathsf{E}\left[\boldsymbol{h}\bar{\mathsf{y}}^{\mathsf{H}}\right] \mathsf{E}\left[\bar{\mathsf{y}}\bar{\mathsf{y}}^{\mathsf{H}}\right]^{-1} \bar{\boldsymbol{y}} \\
&= \boldsymbol{C}_{h,h}\bar{\boldsymbol{X}}^{\mathsf{H}}\left(\bar{\boldsymbol{X}}\boldsymbol{C}_{h,h}\bar{\boldsymbol{X}}^{\mathsf{H}} + \boldsymbol{C}_{\bar{\eta},\bar{\eta}}\right)^{-1} \bar{\boldsymbol{y}} \\
&= \left(\boldsymbol{C}_{h,h}\bar{\boldsymbol{X}}^{\mathsf{H}}\boldsymbol{C}_{\bar{\eta},\bar{\eta}}{}^{-1}\bar{\boldsymbol{X}} + \boldsymbol{I}_N\right)^{-1} \boldsymbol{C}_{h,h}\bar{\boldsymbol{X}}^{\mathsf{H}}\boldsymbol{C}_{\bar{\eta},\bar{\eta}}{}^{-1}\bar{\boldsymbol{y}},
\end{aligned}$$

i. e.

$$\boxed{\hat{\boldsymbol{h}}_{\mathsf{WF}} = \left(\left(\bar{\boldsymbol{X}}^{\mathsf{H}}\boldsymbol{C}_{\bar{\eta},\bar{\eta}}{}^{-1}\bar{\boldsymbol{X}}\right)^{-1}\boldsymbol{C}_{h,h}^{-1} + \boldsymbol{I}_N\right)^{-1} \hat{\boldsymbol{h}}_{\mathsf{ML}}.} \tag{4.9}$$

[5]The derivation is based on the Matrix Inversion Lemma given by $(\boldsymbol{A}+\boldsymbol{BCD})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}\left(\boldsymbol{C}^{-1}+\boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1}\boldsymbol{D}\boldsymbol{A}^{-1}$.

4.3 Spectral Decomposition

In the following we assume $\boldsymbol{C}_{\bar{\eta},\bar{\eta}} = \sigma_\eta^2 \boldsymbol{I}_{MN}$, $\bar{\boldsymbol{X}}^{\mathsf{H}} \bar{\boldsymbol{X}} = M\sigma_x^2 \boldsymbol{I}_{MN}$, and the spectral decomposition[6]

$$\boxed{\boldsymbol{C}_{h,h} = \sum_{i=1}^{N} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i^{\mathsf{H}} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathsf{H}}.} \tag{4.10}$$

Consequently, the ML (LS) estimator and the MMSE estimator can be written as

$$\hat{\boldsymbol{h}}_{\mathsf{ML}} = \frac{1}{M\sigma_x^2} \bar{\boldsymbol{X}}^{\mathsf{H}} \bar{\boldsymbol{y}} = \frac{1}{M\sigma_x^2} \sum_{i=1}^{M} x_i^* \boldsymbol{y}_i \tag{4.11}$$

$$\hat{\boldsymbol{h}}_{\mathsf{WF}} = \boldsymbol{U} \left(\frac{1}{M\gamma} \boldsymbol{\Lambda}^{-1} + \boldsymbol{I}_N \right)^{-1} \boldsymbol{U}^{\mathsf{H}} \hat{\boldsymbol{h}}_{\mathsf{ML}}, \tag{4.12}$$

where $\gamma = \frac{\sigma_x^2}{\sigma_\eta^2}$ denotes the Signal-to-Noise Ratio (SNR) and

$$\boxed{\hat{\boldsymbol{h}}_{\mathsf{WF}} = \sum_{i=1}^{N} \left(1 + \frac{1}{M\gamma\lambda_i}\right)^{-1} \boldsymbol{u}_i \left(\boldsymbol{u}_i^{\mathsf{H}} \hat{\boldsymbol{h}}_{\mathsf{ML}} \right), \qquad \gamma = \frac{\sigma_x^2}{\sigma_\eta^2}.} \tag{4.13}$$

[6] $\boldsymbol{U} = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N] \in \mathbb{C}^{N\times N}$ and $\boldsymbol{\Lambda} = \mathrm{diag}\,[\lambda_1, \ldots, \lambda_N]$ with $\boldsymbol{U}^{\mathsf{H}}\boldsymbol{U} = \boldsymbol{I}_N$ and $\lambda_i \geq 0$ for all $i = 1, \ldots, N$.

Interpretation

$$\boxed{\hat{\boldsymbol{h}}_{\mathsf{WF}} = \sum_{i=1}^{N} \underbrace{\left(1 + \frac{1}{M\gamma\lambda_i}\right)^{-1}}_{\text{Weighting}} \underbrace{\boldsymbol{u}_i \overbrace{\left(\boldsymbol{u}_i^{\mathsf{H}} \hat{\boldsymbol{h}}_{\mathsf{ML}}\right)}^{\hat{h}_{\mathsf{ML},i}}}_{\text{Projection onto } \boldsymbol{u}_i}, \qquad \gamma = \frac{\sigma_x^2}{\sigma_\eta^2}.} \tag{4.14}$$

The ML solution can be decomposed into a linear combination of eigenvectors of the channel covariance matrix. Given this decomposition the MSE solution differs from the ML solution by an individual weighting of the ML components.

- λ_i constant, $\gamma \gg 1$ or $M \gg 1$ ($\rightarrow$ scaled Maximum Likelihood, LS):

$$\hat{\boldsymbol{h}}_{\mathsf{WF}} \propto \sum_{i=1}^{N} \boldsymbol{u}_i \left(\boldsymbol{u}_i^{\mathsf{H}} \hat{\boldsymbol{h}}_{\mathsf{ML}}\right) = \boldsymbol{h}_{\mathsf{ML}}. \tag{4.15}$$

- $\gamma \rightarrow 0$ ($\rightarrow$ Matched Filter):

$$\hat{\boldsymbol{h}}_{\mathsf{WF}} \rightarrow M\gamma \sum_{i=1}^{N} \lambda_i \boldsymbol{u}_i \left(\boldsymbol{u}_i^{\mathsf{H}} \hat{\boldsymbol{h}}_{\mathsf{ML}}\right) \propto \boldsymbol{C}_{h,h} \boldsymbol{h}_{\mathsf{ML}}. \tag{4.16}$$

4.4 Bias and Variance

The mean square error of an estimator can be divided into two error sources, the squared bias and the variance:

$$\boxed{\mathsf{E}\left[\left\|\boldsymbol{h}-\hat{\boldsymbol{h}}\right\|_2^2 \mid \boldsymbol{h}=\boldsymbol{h}\right] = \left[\mathsf{Bias}(\hat{\boldsymbol{h}} \mid \boldsymbol{h}=\boldsymbol{h})\right]^2 + \mathsf{Var}(\hat{\boldsymbol{h}} \mid \boldsymbol{h}=\boldsymbol{h}).} \tag{4.17}$$

- Estimation Bias:

$$\left\|\mathsf{Bias}(\hat{\boldsymbol{h}} \mid \boldsymbol{h}=\boldsymbol{h})\right\|^2 = \left\|\mathsf{E}\left[\hat{\boldsymbol{h}}\right]-\boldsymbol{h}\right\|_2^2 \tag{4.18}$$

- Estimation Variance:

$$\mathsf{Var}(\hat{\boldsymbol{h}} \mid \boldsymbol{h}=\boldsymbol{h}) = \mathsf{E}\left[\left\|\hat{\boldsymbol{h}}-\mathsf{E}\left[\hat{\boldsymbol{h}}\right]\right\|_2^2\right] \tag{4.19}$$

The averaged mean square error is given by

$$\boxed{\mathsf{E}\left[\mathsf{E}\left[\left\|\boldsymbol{h}-\hat{\boldsymbol{h}}\right\|_2^2 \mid \boldsymbol{h}\right]\right] = \mathsf{E}\left[\left\|\boldsymbol{h}-\hat{\boldsymbol{h}}\right\|_2^2\right].} \tag{4.20}$$

- ML (LS) Estimation:

$$\boxed{\left\|\mathsf{Bias}(\hat{\boldsymbol{h}}_{\mathsf{ML}} \mid \boldsymbol{h} = \boldsymbol{h})\right\|^2 = 0,} \tag{4.21}$$

$$\boxed{\mathsf{Var}(\hat{\boldsymbol{h}}_{\mathsf{ML}} \mid \boldsymbol{h} = \boldsymbol{h}) = \frac{N\sigma_\eta^2}{M\sigma_x^2}.} \tag{4.22}$$

- MSE Estimation:

$$\boxed{\left\|\mathsf{Bias}(\hat{\boldsymbol{h}}_{\mathsf{WF}} \mid \boldsymbol{h} = \boldsymbol{h})\right\|^2 = \sum_{i=1}^{N} \left(\frac{1}{1+M\gamma\lambda_i}\right)^2 |h_i|^2,} \tag{4.23}$$

$$\boxed{\mathsf{Var}(\hat{\boldsymbol{h}}_{\mathsf{WF}} \mid \boldsymbol{h} = \boldsymbol{h}) = \sum_{i=1}^{N} \left(\frac{M\gamma\lambda_i}{1+M\gamma\lambda_i}\right)^2 \frac{1}{M\gamma}.} \tag{4.24}$$

Note that the MSE estimator's bias depends on the instantaneous realization of $\boldsymbol{h}$ by means of

$$h_i = \boldsymbol{u}_i^{\mathsf{H}} \boldsymbol{h}, \tag{4.25}$$

whereas the variance of both estimators is independent of the realization $\boldsymbol{h}$.

4.5 Averaged MSE

The averaged mean square error of the two estimators is

$$\boxed{\mathsf{MSE}(\hat{\boldsymbol{h}}_{\mathsf{ML}}) = \frac{N\sigma_\eta^2}{M\sigma_{\mathsf{x}}^2} = \frac{N}{M\gamma},} \tag{4.26}$$

where $\gamma = \frac{\sigma_{\mathsf{x}}^2}{\sigma_\eta^2}$ denotes the Signal-to-Noise Ratio (SNR) and

$$\begin{aligned}
\mathsf{MSE}(\hat{\boldsymbol{h}}_{\mathsf{WF}}) &= \mathsf{E}\left[\left[\mathsf{Bias}(\hat{\boldsymbol{h}} \mid \boldsymbol{h} = \boldsymbol{h})\right]^2 + \mathsf{Var}(\hat{\boldsymbol{h}} \mid \boldsymbol{h} = \boldsymbol{h})\right] \\
&= \mathsf{E}\left[\sum_{i=1}^{N}\left(\frac{1}{1+M\gamma\lambda_i}\right)^2 |h_i|^2 + \sum_{i=1}^{N}\left(\frac{M\gamma\lambda_i}{1+M\gamma\lambda_i}\right)^2 \frac{1}{M\gamma}\right] \\
&= \mathsf{E}\left[\sum_{i=1}^{N}\lambda_i\frac{\frac{|h_i|^2}{\lambda_i}+M\gamma\lambda_i}{(1+M\gamma\lambda_i)^2}\right] = \sum_{i=1}^{N}\lambda_i\frac{\frac{\mathsf{E}\left[|h_i|^2\right]}{\lambda_i}+M\gamma\lambda_i}{(1+M\gamma\lambda_i)^2},
\end{aligned}$$

i. e.

$$\boxed{\mathsf{MSE}(\hat{\boldsymbol{h}}_{\mathsf{WF}}) = \sum_{i=1}^{N}\frac{\lambda_i}{1+M\gamma\lambda_i}.} \tag{4.27}$$

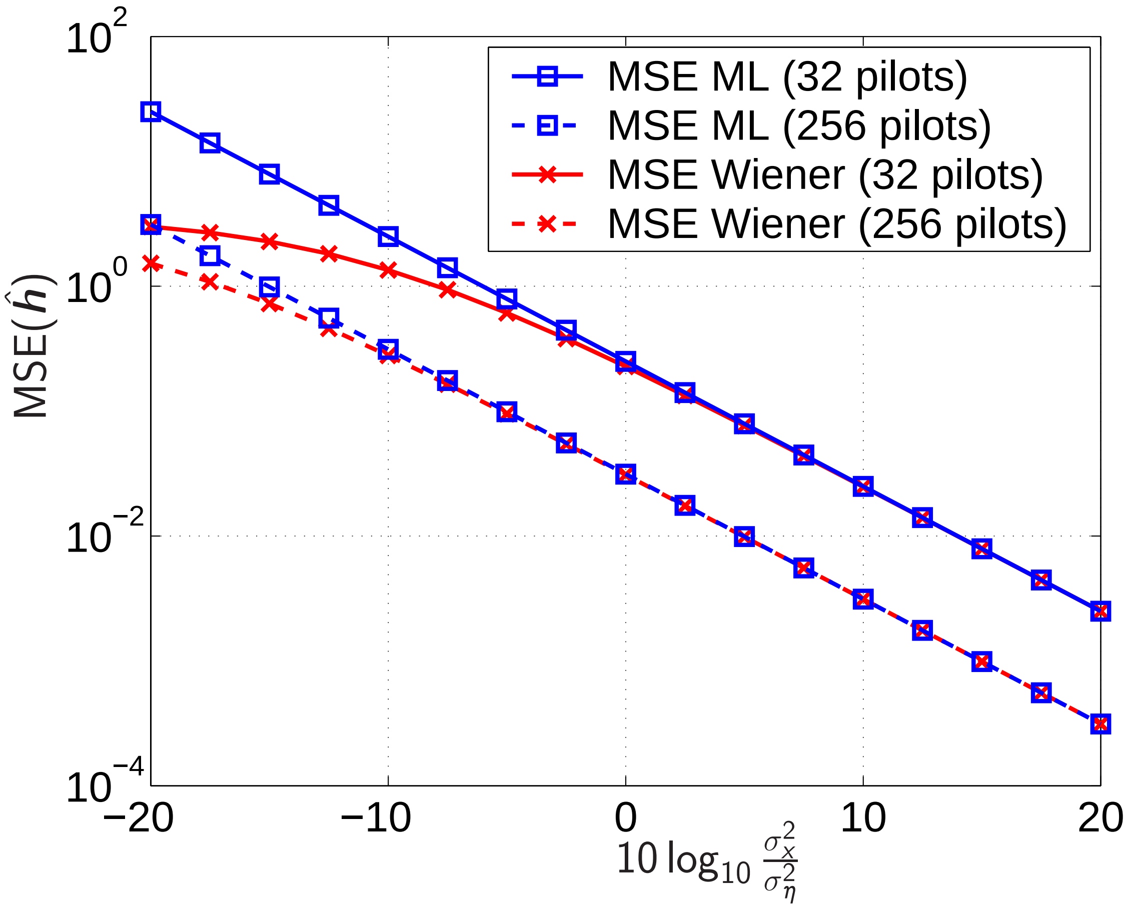

Fig. 4.1. Comparison of ML and MSE Channel Estimation in Terms of Mean Square Error: $\boldsymbol{h} \in \mathbb{C}^N$, $N = 8$, $\sigma_{h,l}^2 = \exp(-\kappa_\ell/8)$, and $\kappa_\ell \in \{0, 16\}$.

5. Pilot Assisted Channel Estimation in OFDM

In OFDM systems either the channel impulse response function (time domain) or the channel transfer function (frequency domain) can be estimated.

The two alternatives differ from each other in terms of numerical complexity and estimation quality.

In order to keep with the OFDM system specifications it seems preferable to estimate the channel transfer function directly but still preserving the advantages of time domain estimation.

The signal model reads

$$\boxed{\boldsymbol{Y}_i = \boldsymbol{X}_i \boldsymbol{S} \boldsymbol{H} + \boldsymbol{\eta}_i} \tag{5.1}$$

where $\boldsymbol{Y}_i \in \mathbb{C}^P$, $\boldsymbol{H} \in \mathbb{C}^N$, $\boldsymbol{\eta}_i \in \mathbb{C}^P$ are vectors,

$$\boldsymbol{X}_i = \mathsf{diag}\left[X_{i,1}, \ldots, X_{i,P}\right] \;\in\; \mathbb{C}^{P \times P}, \tag{5.2}$$

and $\boldsymbol{S} \in \mathbb{C}^{P \times N}$ denotes the matrix which selects P-out-of-N channel coefficients:

$$\boldsymbol{S} : \mathbb{C}^N \to \mathbb{C}^P,\; \boldsymbol{H} \mapsto \boldsymbol{H}_P = [H_{s_1}, \ldots, H_{s_P}]^\mathsf{T}. \tag{5.3}$$

- Since in general time-variant channels are considered, the number of transmitted pilot vectors (OFDM symbols) is limited to $M = 1$, thus

$$\boldsymbol{Y}_i, \boldsymbol{X}_i, \boldsymbol{\eta}_i \quad \rightarrow \quad \boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{\eta}.$$

- If not all tones (OFDM carriers) are used for channel estimation, i. e. $P < N$ and $\boldsymbol{S} \neq \boldsymbol{I}_N$ respectively, a subsequent interpolation from $\boldsymbol{H}_P \rightarrow \boldsymbol{H}$ must be carried out.
- An interpolation requires an underlying relation between the channel coefficients of adjacent tones.
- In general there is a strong dependency between the H_i which depends on the multipath characteristic of the corresponding channel impulse response function of the OFDM channel.
- Estimation and interpolation can be carried out in two alternative directions:
  - Estimation of $\boldsymbol{H}_P$ and subsequent interpolation of $\boldsymbol{H}$.
  - Joint estimation and interpolation of $\boldsymbol{H}$ by exploiting the properties of the channel impulse response function $\boldsymbol{h}$.

5.1 Channel Estimation for Fully Loaded OFDM Systems ($P = N$)

- LS estimation based on $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{H} + \boldsymbol{\eta}$ and $\boldsymbol{X}^{\mathsf{H}}\boldsymbol{X} = \sigma_{\mathsf{x}}^2 \boldsymbol{I}_N$:

$$\boxed{\hat{\boldsymbol{H}}_{\mathsf{LS}} = \boldsymbol{X}^{-1}\boldsymbol{Y},} \tag{5.4}$$

with

$$\mathsf{MSE}\left(\hat{H}_{\mathsf{LS}}\right) = \mathsf{E}\left[\mathsf{tr}\left[\boldsymbol{X}^{-1}\boldsymbol{\eta}\boldsymbol{\eta}\boldsymbol{X}^{-\mathsf{H}}\right]\right] = \frac{N}{\gamma}. \tag{5.5}$$

- MSE estimation based on $\boldsymbol{Y}$, $\boldsymbol{X}^{\mathsf{H}}\boldsymbol{X} = \sigma_{\mathsf{x}}^2 \boldsymbol{I}_N$, and prior knowledge $\hat{\boldsymbol{C}}_{H,H}$ about the channel correlation matrix $\boldsymbol{C}_{H,H}$:

$$\boxed{\hat{\boldsymbol{H}}_{\mathsf{WF}} = \hat{\boldsymbol{C}}_{H,H}\left(\hat{\boldsymbol{C}}_{H,H} + \tfrac{1}{\gamma}\boldsymbol{I}_N\right)^{-1}\hat{\boldsymbol{H}}_{\mathsf{LS}},} \tag{5.6}$$

with

$$\mathsf{MSE}\left(\hat{H}_{\mathsf{WF}}\right) = \sum_{i=1}^{N} \frac{\frac{\zeta_i}{\hat{\lambda}_i} + \gamma\hat{\lambda}_i}{1+\gamma\hat{\lambda}_i} \cdot \frac{\hat{\lambda}_i}{1+\gamma\hat{\lambda}_i}, \tag{5.7}$$

where the $\hat{\lambda}_i$ are the eigenvalues of $\hat{\boldsymbol{C}}_{H,H} = \hat{\boldsymbol{U}}\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{U}}^{\mathsf{H}}$ and ζ_i are the main diagonal elements of $\hat{\boldsymbol{U}}^{\mathsf{H}}\boldsymbol{C}_{H,H}\hat{\boldsymbol{U}}$.

5.2 Reduced-Rank Channel Estimation for Fully Loaded OFDM Systems

Given prior knowledge about the channel impulse response function of the transmission channel, i. e.

$$h(t) = \sum_{\ell=1}^{L} h_\ell \, \delta(t - \kappa_\ell T_\mathsf{a}), \qquad \kappa_\ell \in \{0, 1, \ldots, N_\mathsf{p}\}, \tag{5.8}$$

the sample vector $\boldsymbol{H}$ of the channel transfer function can be constrained on a L-dimensional subspace:

$$\boxed{\boldsymbol{H} \ \in \ \mathsf{span} \underbrace{\left[\boldsymbol{f}_{\kappa_1+1}, \boldsymbol{f}_{\kappa_2+1}, \ldots, \boldsymbol{f}_{\kappa_L+1}\right]}_{\boldsymbol{F}_L \in \mathbb{C}^{N \times L}},} \tag{5.9}$$

where $\boldsymbol{f}_i = \boldsymbol{F}\boldsymbol{e}_i = \left[1, \mathsf{e}^{\mathsf{j}\frac{2\pi}{N}(i-1)}, \ldots, \mathsf{e}^{\mathsf{j}\frac{2\pi}{N}(N-1)(i-1)}\right]^\mathsf{T}$.

If the CIR function is unknown, given the assumption $\max\limits_{\ell=1,\ldots,L} \kappa_\ell \leq N_p$, the channel vector can still be constrained on

$$\boldsymbol{F}_{N_\mathsf{p}+1} = \mathsf{span}\left[\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_{N_\mathsf{p}+1}\right]. \tag{5.10}$$

LS Estimation

$$\begin{aligned}\hat{\boldsymbol{H}}_{\mathsf{LS}}^{(\mathsf{red})} &= \underset{\boldsymbol{H}}{\mathsf{argmin}} \left\| \boldsymbol{Y} - \boldsymbol{X}\boldsymbol{H} \right\|^2_{\boldsymbol{C}^{-1}_{\eta,\eta}} \quad \text{s. t.} \quad \boldsymbol{H} = \boldsymbol{F}_L \boldsymbol{h}, \quad \boldsymbol{h} \in \mathbb{C}^L \\ &= \boldsymbol{F}_L \left(\boldsymbol{F}_L^{\mathsf{H}} \boldsymbol{X}^{\mathsf{H}} \boldsymbol{X} \boldsymbol{F}_L \right)^{-1} \boldsymbol{F}_L^{\mathsf{H}} \boldsymbol{X}^{\mathsf{H}} \boldsymbol{Y}, \end{aligned} \tag{5.11}$$

i. e. the unbiased estimator[1] reads

$$\boxed{\hat{\boldsymbol{H}}_{\mathsf{LS}}^{(\mathsf{red})} = \boldsymbol{F}_L \left(\boldsymbol{F}_L^{\mathsf{H}} \boldsymbol{F}_L \right)^{-1} \boldsymbol{F}_L^{\mathsf{H}} \hat{\boldsymbol{H}}_{\mathsf{LS}} = \underbrace{\frac{1}{N} \boldsymbol{F}_L \boldsymbol{F}_L^{\mathsf{H}}}_{\text{Projection}} \hat{\boldsymbol{H}}_{\mathsf{LS}}.} \tag{5.12}$$

The MSE of the reduced-rank channel estimator equals

$$\mathsf{MSE}\left(\hat{H}_{\mathsf{LS}}^{(\mathsf{red})} \right) = \frac{1}{N^2} \mathsf{E} \left[\mathsf{tr} \left[\boldsymbol{F}_L \boldsymbol{F}_L^{\mathsf{H}} \boldsymbol{X}^{-1} \boldsymbol{\eta} \boldsymbol{\eta} \boldsymbol{X}^{-\mathsf{H}} \boldsymbol{F}_L \boldsymbol{F}_L^{\mathsf{H}} \right] \right] = \frac{L}{\gamma}. \tag{5.13}$$

If the CIR function is unknown, the projection can still be based on $\mathsf{span}\left[\boldsymbol{F}_{N_{\mathsf{p}}+1} \right]$ and the MSE is equal to[2]

$$\frac{L}{\gamma} \leq \mathsf{MSE}\left(\hat{H}_{\mathsf{LS}}^{(\mathsf{red})} \right) = \frac{N_{\mathsf{p}} + 1}{\gamma} \ll \frac{N}{\gamma}. \tag{5.14}$$

[1] If the subspace assumption of the channel vector holds, the reduced-rank LS estimator is still unbiased.

[2] $\boldsymbol{F}_{N_{\mathsf{p}}+1} = [\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{f}_{N_{\mathsf{p}}+1}]$.

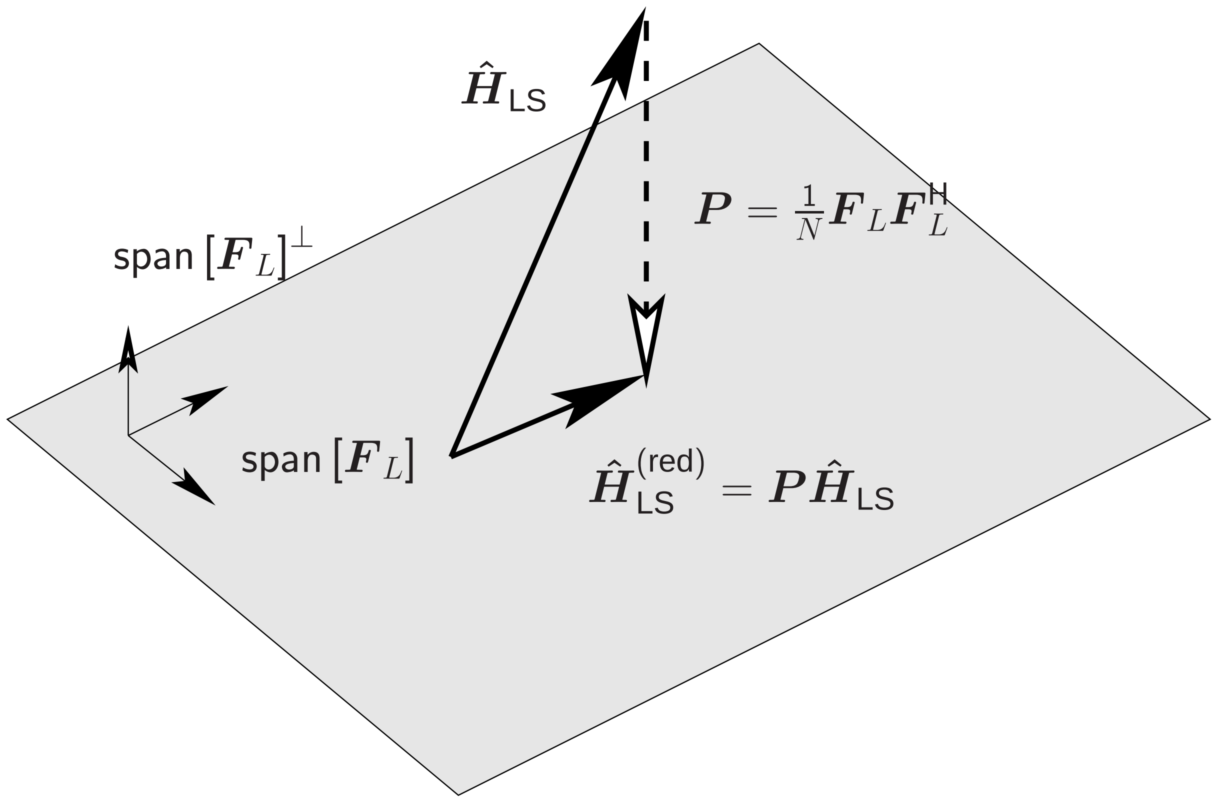

Fig. 5.1. Projection of the LS Estimate to the Reduced-Rank Subspace of the Channel Vector.

MMSE Estimation

Given the constraint $\boldsymbol{H} = \boldsymbol{F}_L \boldsymbol{h}$, the MSE estimator reads

$$\boxed{\hat{\boldsymbol{H}}_{\mathsf{WF}}^{(\mathsf{red})} = \boldsymbol{F}_L \boldsymbol{R}_{h,h} \boldsymbol{F}_L^{\mathsf{H}} \left(\boldsymbol{F}_L \boldsymbol{R}_{h,h} \boldsymbol{F}_L^{\mathsf{H}} + \tfrac{1}{\gamma} \boldsymbol{I}_N \right)^{-1} \hat{\boldsymbol{H}}_{\mathsf{LS}},} \tag{5.15}$$

where $\boldsymbol{R}_{h,h}$ denotes the correlation matrix of the L channel coefficients h_ℓ from the CIR function.

If the CIR function is unknown, we assume the h_ℓ to be uncorrelated with uniform power delay profile, i.e. $\hat{\boldsymbol{R}}_{h,h} = \hat{\sigma}_h^2 \boldsymbol{I}_L$. Then, the MSE estimator can be obtained as

$$\hat{\hat{\boldsymbol{H}}}_{\mathsf{WF}}^{(\mathsf{red})} = \boldsymbol{F}_L \boldsymbol{F}_L^{\mathsf{H}} \left(\boldsymbol{F}_L \boldsymbol{F}_L^{\mathsf{H}} + \frac{1}{\gamma \hat{\sigma}_h^2} \boldsymbol{I}_N \right)^{-1} \hat{\boldsymbol{H}}_{\mathsf{LS}}, \tag{5.16}$$

with the averaged

$$\mathsf{MSE}\left(\hat{\hat{H}}_{\mathsf{WF}}^{(\mathsf{red})} \right) = \sum_{i=1}^{L} \frac{\frac{\sigma_{h,i}^2}{\hat{\sigma}_h^2} + \gamma N \hat{\sigma}_h^2}{1 + \gamma N \hat{\sigma}_h^2} \cdot \frac{N \hat{\sigma}_h^2}{1 + \gamma N \hat{\sigma}_h^2} \approx \frac{L}{\frac{1}{N \hat{\sigma}_h^2} + \gamma}, \tag{5.17}$$

where the $\sigma_{h,i}^2$ are the main diagonal elements of the true $\boldsymbol{C}_{h,h}$.

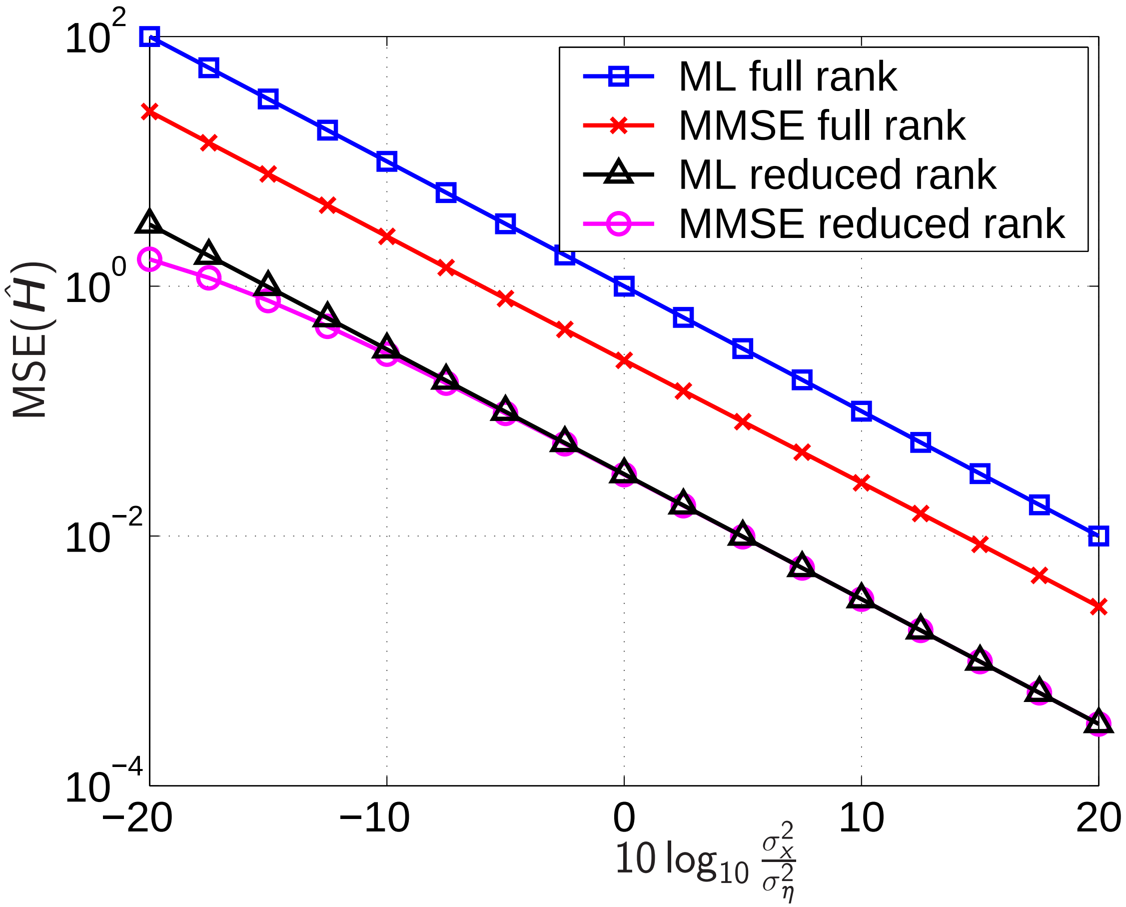

Fig. 5.2. Comparison of (Reduced-Rank) Channel Estimation in Fully Loaded OFDM Systems: $N = 256$, $L = 8$, $\sigma_{h,l}^2 = \exp(-\kappa_\ell/8)$, and $\kappa_\ell \in \{0, 16\}$

5.3 Channel Estimation for Non-Fully Loaded OFDM Systems ($P < N$)

If not all OFDM tones are used for channel estimation, after the estimation of $\boldsymbol{H}_P = [H_{s_1}, \ldots, H_{s_P}]^\mathsf{T}$ the unknown elements of $\boldsymbol{H}$ must be interpolated from the estimate $\boldsymbol{H}_P$.

Alternatively, a joint estimation and interpolation of $\boldsymbol{H}$ by means of the incomplete observations $\boldsymbol{Y}_P$ and exploiting the properties of the CIR function can be carried out.

LS Channel Estimation and Interpolation

Given an estimate of $\boldsymbol{H}_P = \boldsymbol{S}\boldsymbol{H}$ based on the observation vector $\boldsymbol{Y}_P$ from $P \geq L$ pilot carriers,[3] the LS interpolation of $\boldsymbol{H}$ can be carried out by means of

- linear and nonlinear Splines,
- kernel functions (e.g. sinc-function), and by
- exploiting the subspace properties of the channel transfer function, viz.

$$\boldsymbol{H} \in \mathsf{span}\left[\boldsymbol{F}_L\right], \qquad \boldsymbol{F}_L = [\boldsymbol{f}_{\kappa_1+1}, \boldsymbol{f}_{\kappa_2+1}, \ldots, \boldsymbol{f}_{\kappa_L+1}].$$

[3] $P \geq N_\mathsf{p}$, if the CIR is unknown.

Given the LS estimate of $\boldsymbol{H}_P$ as

$$\hat{\boldsymbol{H}}_{P,\mathsf{LS}} = \boldsymbol{X}_P^{-1}\boldsymbol{Y}_P. \tag{5.18}$$

and the subspace constraint

$$\boxed{\boldsymbol{H} \in \mathsf{span}\left[\boldsymbol{F}_L\right] \quad \Rightarrow \quad \boldsymbol{H}_P \in \mathsf{span}\left[\boldsymbol{F}_{P\times L}\right],} \tag{5.19}$$

where[4]

$$\boldsymbol{F}_{P\times L} = \boldsymbol{S}\boldsymbol{F}_L = \begin{bmatrix} \boldsymbol{e}_{s_1}^{\mathsf{T}} \\ \boldsymbol{e}_{s_2}^{\mathsf{T}} \\ \vdots \\ \boldsymbol{e}_{s_P}^{\mathsf{T}} \end{bmatrix} \boldsymbol{F}\left[\boldsymbol{e}_{\kappa_1}\ \boldsymbol{e}_{\kappa_2}\ \cdots\ \boldsymbol{e}_{\kappa_L}\right],$$

the LS interpolation of $\boldsymbol{H} = \boldsymbol{F}_L\boldsymbol{h}$ by means of $\boldsymbol{H}_P = \boldsymbol{S}\boldsymbol{H} = \boldsymbol{S}\boldsymbol{F}_L\boldsymbol{h} = \boldsymbol{F}_{P\times L}\boldsymbol{h}$ can be obained as

$$\boxed{\hat{\boldsymbol{H}}_{\mathsf{LS}}^{(\mathsf{int})} = \boldsymbol{F}_L\left(\boldsymbol{F}_{P\times L}^{\mathsf{H}}\boldsymbol{F}_{P\times L}\right)^{-1}\boldsymbol{F}_{P\times L}^{\mathsf{H}}\hat{\boldsymbol{H}}_{P,\mathsf{LS}}.} \tag{5.20}$$

[4] $[\boldsymbol{F}_{P\times L}]_{i,j} = \boldsymbol{e}_{s_i}^{\mathsf{T}}\boldsymbol{F}\boldsymbol{e}_{\kappa_j}$, where the i-th unity vector $\boldsymbol{e}_i$ denotes the i-th column vector of the unity matrix $\boldsymbol{I}_N$.

The averaged MSE of the LS interpolation of $\boldsymbol{H}$ is given by[5]

$$\begin{aligned}
\mathsf{MSE}\left(\hat{H}_{\mathsf{LS}}^{(\mathsf{int})}\right) &= \sigma_\eta^2 \cdot \mathsf{tr}\left[\boldsymbol{F}_L \boldsymbol{F}_{P\times L}^{+}\left(\boldsymbol{X}_P^{\mathsf{H}}\boldsymbol{X}_P\right)^{-1}\boldsymbol{F}_{P\times L}^{+,\mathsf{H}}\boldsymbol{F}_L^{\mathsf{H}}\right] && (5.21)\\
&= \sigma_\eta^2 N \sum_{i=1}^{P} \frac{\left\|\boldsymbol{F}_{P\times L}^{+}\boldsymbol{e}_i\right\|_2^2}{|X_{P,i}|^2} && (5.22)\\
&\geq \sigma_\eta^2 P N \left(\prod_{i=1}^{P} \frac{\left\|\boldsymbol{F}_{P\times L}^{+}\boldsymbol{e}_i\right\|_2^2}{|X_{P,i}|^2}\right)^{\frac{1}{P}}, && (5.23)
\end{aligned}$$

with equality if and only if $\left\|\boldsymbol{F}_{P\times L}^{+}\boldsymbol{e}_i\right\|_2^2 / |X_{P,i}|^2$ is equally distributed.

Given the constraint $\mathsf{tr}\left[\boldsymbol{X}_P^{\mathsf{H}}\boldsymbol{X}_P\right] = \sigma_x^2 P$ and a **fixed** selection matrix $\boldsymbol{S}$, the optimization of $\boldsymbol{X}_P$ based on (5.23) leads to[6]

$$\boxed{\left|X_{P,i}^{(\mathsf{opt})}\right|^2 = \sigma_x^2 P \frac{\left\|\boldsymbol{F}_{P\times L}^{+}\boldsymbol{e}_i\right\|_2^2}{\sum_{i=1}^{P}\left\|\boldsymbol{F}_{P\times L}^{+}\boldsymbol{e}_i\right\|_2^2} = \sigma_x^2 P \frac{\left\|\boldsymbol{F}_{P\times L}^{+}\boldsymbol{e}_i\right\|_2^2}{\left\|\boldsymbol{F}_{P\times L}^{+}\right\|_F^2}.} \quad (5.24)$$

[5] $\boldsymbol{F}_{P\times L}^{+} = \left(\boldsymbol{F}_{P\times L}^{\mathsf{H}}\boldsymbol{F}_{P\times L}\right)^{-1}\boldsymbol{F}_{P\times L}^{\mathsf{H}}$ and $\boldsymbol{F}_{P\times L} = \boldsymbol{S}\boldsymbol{F}_L$.

[6] The solution is not unique.

Finally, the resulting MSE reads[7]

$$\begin{aligned}\mathsf{MSE}\left(\hat{H}_{\mathsf{LS}}^{(\mathsf{int})}, \boldsymbol{X}_P^{(\mathsf{opt})}\right) &= \frac{N}{\gamma}\left\|\boldsymbol{F}_{P\times L}^{+}\right\|_F^2 = \frac{N}{\gamma}\,\mathsf{tr}\left[\left(\boldsymbol{F}_{P\times L}^{\mathsf{H}}\boldsymbol{F}_{P\times L}\right)^{-1}\right] &(5.25)\\ &\geq \frac{N}{\gamma}L^2\left(\mathsf{tr}\left[\boldsymbol{F}_{P\times L}^{\mathsf{H}}\boldsymbol{F}_{P\times L}\right]\right)^{-1} = \frac{NL}{P\gamma}, &(5.26)\end{aligned}$$

with equality if all eigenvalues of $\boldsymbol{F}_{P\times L}^{\mathsf{H}}\boldsymbol{F}_{P\times L}$ are identical.

In order to minimize the resulting MSE the P rows of $\boldsymbol{F}_{P\times L}$ must be selected from $\boldsymbol{F}_L$ such that the columns of $\boldsymbol{F}_{P\times L}$ are orthogonal to each other, i. e. an appropriate choice (due to the self-similarity properties of the DFT) is an equal distance selection of pilot carriers:[8]

$$\boxed{\forall i = 1, \ldots, P: \quad s_i = \left(i - \tfrac{1}{2}\right)\tfrac{N}{P}, \quad \log_2 \tfrac{N}{P} \in \mathbb{N},\ P \geq L.} \tag{5.27}$$

The resulting MSE is

$$\boxed{\mathsf{MSE}\left(\hat{H}_{\mathsf{LS}}^{(\mathsf{int})}, \boldsymbol{X}_P^{(\mathsf{opt})}, \boldsymbol{S}^{(\mathsf{opt})}\right) = \frac{NL}{P\gamma}.} \tag{5.28}$$

[7] $\left\|\boldsymbol{F}_{P\times L}^{+}\right\|_F^2 = \mathsf{tr}\left[\boldsymbol{F}_{P\times L}^{+,\mathsf{H}}\boldsymbol{F}_{P\times L}^{+}\right] = \mathsf{tr}\left[\boldsymbol{F}_{P\times L}\left(\boldsymbol{F}_{P\times L}^{\mathsf{H}}\boldsymbol{F}_{P\times L}\right)^{-2}\boldsymbol{F}_{P\times L}^{\mathsf{H}}\right] = \mathsf{tr}\left[\left(\boldsymbol{F}_{P\times L}^{\mathsf{H}}\boldsymbol{F}_{P\times L}\right)^{-1}\right]$.

[8] This confirms the optimum choice of $\boldsymbol{X}_P = \sigma_{\mathsf{x}}^2\boldsymbol{I}_P$ as in fully-loaded OFDM systems.

Joint LS Estimation and Interpolation

Given the subspace constrained signal model $\boldsymbol{Y}_P = \boldsymbol{X}_P \boldsymbol{F}_{P\times L}\boldsymbol{h} + \boldsymbol{\eta}_P$ behind the observation vector, the joint LS estimation & interpolation of $\boldsymbol{H}$ can be determined to

$$\hat{\boldsymbol{H}}_{\mathsf{LS}}^{(\mathsf{joint})} = \boldsymbol{F}_L \left(\boldsymbol{F}_{P\times L}^{\mathsf{H}} \boldsymbol{X}_P^{\mathsf{H}} \boldsymbol{X}_P \boldsymbol{F}_{P\times L}\right)^{-1} \boldsymbol{F}_{P\times L}^{\mathsf{H}} \boldsymbol{X}_P^{\mathsf{H}} \boldsymbol{Y}_P. \tag{5.29}$$

The averaged MSE of the LS estimation & interpolation of $\boldsymbol{H}$ is given as

$$\begin{aligned}
\mathsf{MSE}\left(\hat{H}_{\mathsf{LS}}^{(\mathsf{joint})}\right) &= \sigma_\eta^2 N \,\mathsf{tr}\left[\left(\boldsymbol{F}_{P\times L}^{\mathsf{H}} \boldsymbol{X}_P^{\mathsf{H}} \boldsymbol{X}_P \boldsymbol{F}_{P\times L}\right)^{-1}\right] && (5.30)\\
&\geq \sigma_\eta^2 N L^2 \left(\mathsf{tr}\left[\boldsymbol{F}_{P\times L}^{\mathsf{H}} \boldsymbol{X}_P^{\mathsf{H}} \boldsymbol{X}_P \boldsymbol{F}_{P\times L}\right]\right)^{-1} && (5.31)\\
&= \sigma_\eta^2 N L^2 \left(\sum_{i=1}^{P} L|X_{P,i}|^2\right)^{-1} = \frac{NL}{P\gamma}, && (5.32)
\end{aligned}$$

with equality if all eigenvalues of $\boldsymbol{F}_{P\times L}^{\mathsf{H}} \boldsymbol{X}_P^{\mathsf{H}} \boldsymbol{X}_P \boldsymbol{F}_{P\times L}$ are identical. This confirms the optimum choice of $\boldsymbol{X}_P = \sigma_{\mathsf{x}}^2 \boldsymbol{I}_P$ in terms of averaged MSE also in non-fully loaded OFDM systems ($P < N$).

MSE Estimation and Interpolation

Given the constraint $\boldsymbol{H} = \boldsymbol{F}_L \boldsymbol{h}$, $\mathrm{tr}\left[\boldsymbol{X}_P^{\mathsf{H}} \boldsymbol{X}_P\right] = \sigma_{\mathsf{x}}^2 P$, and $\boldsymbol{F}_{P\times L}$, the MSE estimation & interpolation can be obtained similar to (5.15) as

$$\boxed{\hat{\boldsymbol{H}}_{\mathsf{WF}}^{(\mathsf{int})} = \boldsymbol{F}_L \hat{\boldsymbol{R}}_{h,h} \boldsymbol{F}_{P\times L}^{\mathsf{H}} \left(\boldsymbol{F}_{P\times L} \hat{\boldsymbol{R}}_{h,h} \boldsymbol{F}_{P\times L}^{\mathsf{H}} + \tfrac{1}{\gamma} \boldsymbol{I}_P\right)^{-1} \hat{\boldsymbol{H}}_{P,\mathsf{LS}}.} \tag{5.33}$$

If the CIR function is unknown, we again assume the h_ℓ to be uncorrelated with uniform power delay profile, i. e. $\hat{\boldsymbol{C}}_{h,h} = \hat{\sigma}_h^2 \boldsymbol{I}_L$. The MSE estimator is equal to

$$\hat{\hat{\boldsymbol{H}}}_{\mathsf{WF}}^{(\mathsf{int})} = \boldsymbol{F}_L \boldsymbol{F}_{P\times L}^{\mathsf{H}} \left(\boldsymbol{F}_{P\times L} \boldsymbol{F}_{P\times L}^{\mathsf{H}} + \frac{1}{\gamma \hat{\sigma}_h^2} \boldsymbol{I}_P\right)^{-1} \hat{\boldsymbol{H}}_{P,\mathsf{LS}}, \tag{5.34}$$

with

$$\mathsf{MSE}\left(\hat{\hat{H}}_{\mathsf{WF}}^{(\mathsf{int})}\right) = \sum_{i=1}^{L} \frac{\frac{\sigma_{h,i}^2}{\hat{\sigma}_h^2} + \gamma P \hat{\sigma}_h^2}{1 + \gamma P \hat{\sigma}_h^2} \cdot \frac{N \hat{\sigma}_h^2}{1 + \gamma P \hat{\sigma}_h^2} \approx \frac{N}{P} \cdot \frac{L}{\frac{1}{P\hat{\sigma}_h^2} + \gamma}, \tag{5.35}$$

where the $\sigma_{h,i}^2$ are the main diagonal elements of the true $\boldsymbol{C}_{h,h}$.

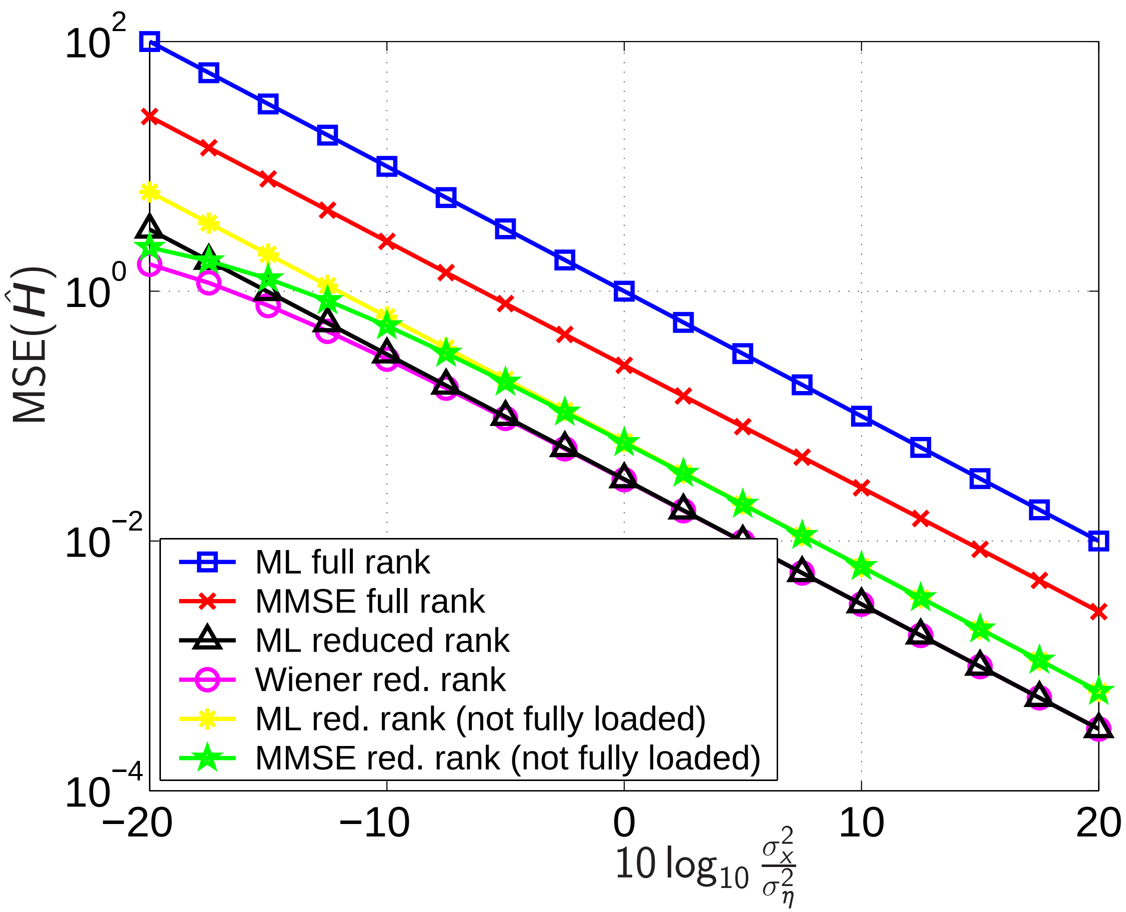

Fig. 5.3. Comparison of (Reduced-Rank) Channel Estimation in Non-Fully Loaded OFDM Systems: $N = 256$, $P = 128$, $L = 8$, $\sigma_{h,l}^2 = \exp(-\kappa_\ell/8)$, and $\kappa_\ell \in \{0, 16\}$

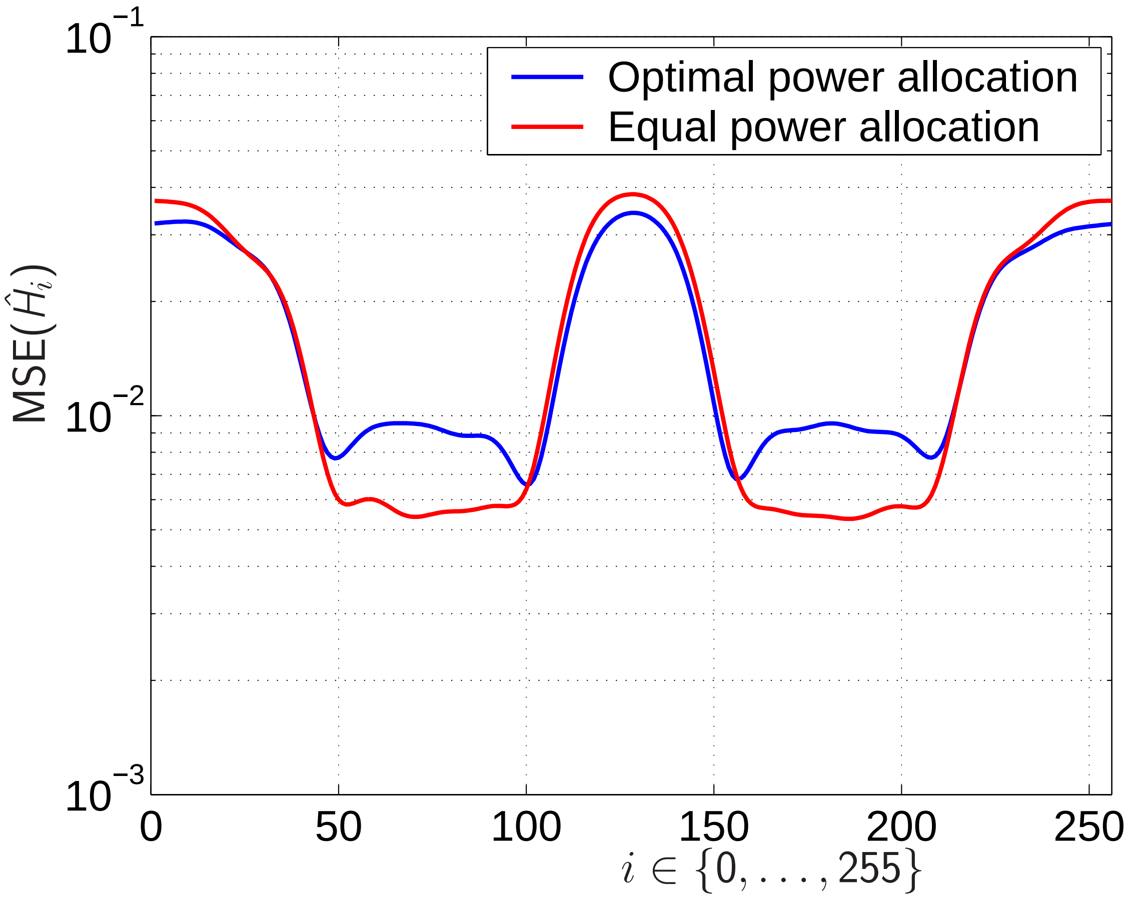

Fig. 5.4. MSE Distribution in Non-Fully Loaded OFDM Systems: $N = 256$, $P = 2 \times 64$ Pilots (**separated in two disjoint blocks of adjacent tones**), $L = 8$, $\sigma^2_{h,l} = \exp(-\kappa_\ell/8)$, and $\kappa_\ell \in \{0, 16\}$

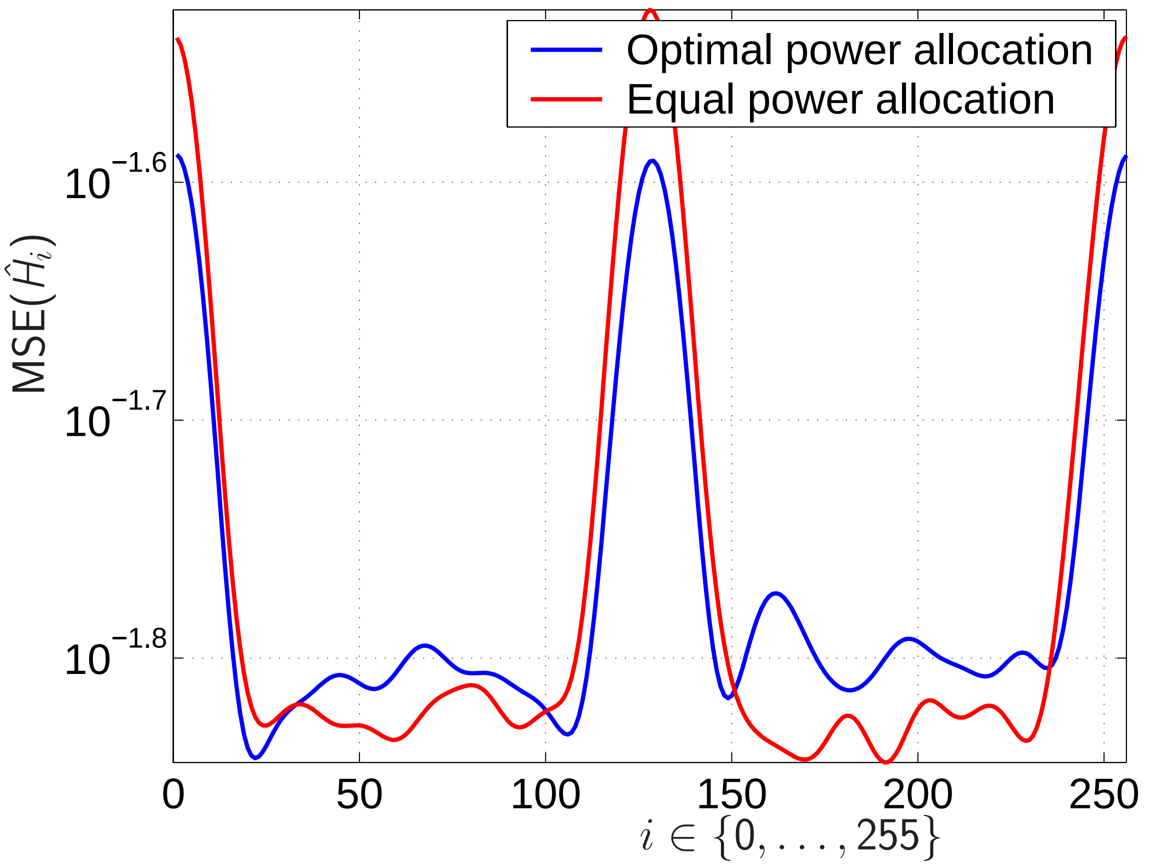

Fig. 5.5. MSE of Subcarrier Channel Estimation in Non-Fully Loaded OFDM Systems: $N = 256$, $P = 2 \times 26$ Pilots (**equally distributed in two disjoint blocks of OFDM tones**), $L = 8$, $\sigma_{h,l}^2 = \exp(-\kappa_\ell/8)$, and $\kappa_\ell \in \{0, 16\}$

6. Extension to Channel Prediction in Time-Variant OFDM

Given a series of observations $\hat{\boldsymbol{H}}_{\mathsf{LS},n}, \hat{\boldsymbol{H}}_{\mathsf{LS},n-1}, \ldots, \hat{\boldsymbol{H}}_{\mathsf{LS},n-M+1}$, the best linear prediction of $\boldsymbol{H}_{n+1}$ is given by

$$\boxed{\hat{\boldsymbol{H}}_{\mathsf{LP}} = \boldsymbol{C}_{H,\bar{Y}} \boldsymbol{C}_{\bar{Y},\bar{Y}}^{-1},} \tag{6.1}$$

where

$$\begin{aligned} \boldsymbol{C}_{H,\bar{Y}} &= \mathsf{E}\left[\boldsymbol{H}_{n+1} \bar{\boldsymbol{Y}}_n^{\mathsf{H}}\right] \\ &= \boldsymbol{F}_L \cdot \mathsf{E}\left[\boldsymbol{h}_{n+1} \bar{\boldsymbol{h}}_n\right] \cdot (\boldsymbol{I}_M \otimes \boldsymbol{F}_L)^{\mathsf{H}} (\boldsymbol{I}_M \otimes \boldsymbol{X})^{\mathsf{H}} \end{aligned} \tag{6.2}$$

and[1]

$$\begin{aligned} \boldsymbol{C}_{\bar{Y},\bar{Y}} &= \mathsf{E}\left[\bar{\boldsymbol{Y}}_n \bar{\boldsymbol{Y}}_n^{\mathsf{H}}\right] \\ &= (\boldsymbol{I}_M \otimes \boldsymbol{X})(\boldsymbol{I}_M \otimes \boldsymbol{F}_L) \cdot \mathsf{E}\left[\bar{\boldsymbol{h}}_n \bar{\boldsymbol{h}}_n^{\mathsf{H}}\right] \cdot (\boldsymbol{I}_M \otimes \boldsymbol{F}_L)^{\mathsf{H}} (\boldsymbol{I}_M \otimes \boldsymbol{X})^{\mathsf{H}} + \sigma_\eta^2 \boldsymbol{I}_{MN}, \end{aligned} \tag{6.3}$$

[1] $\boldsymbol{Y}_{n-i} = \boldsymbol{X}\boldsymbol{H}_{n-i} + \boldsymbol{\eta}_{n-i}$.

with

$$\bar{\boldsymbol{Y}}_n = \left[\boldsymbol{Y}_n^{\mathsf{T}}, \boldsymbol{Y}_{n-1}^{\mathsf{T}}, \cdots, \boldsymbol{Y}_{n-M+1}^{\mathsf{T}}\right]^{\mathsf{T}}$$
$$\bar{\boldsymbol{h}}_n = \left[\boldsymbol{h}_n^{\mathsf{T}}, \boldsymbol{h}_{n-1}^{\mathsf{T}}, \cdots, \boldsymbol{h}_{n-M+1}^{\mathsf{T}}\right]^{\mathsf{T}}.$$

Remark

$$\begin{aligned} \mathsf{E}\left[\boldsymbol{h}_{n+1}\bar{\boldsymbol{h}}_n\right] &= \mathsf{E}\left[\boldsymbol{h}_{n+1}\left[\boldsymbol{h}_n^{\mathsf{H}}, \boldsymbol{h}_{n-1}^{\mathsf{H}}, \cdots, \boldsymbol{h}_{n-M+1}^{\mathsf{H}}\right]\right] \\ &= \underbrace{\left[r_1, r_2, \cdots, r_M\right]}_{\boldsymbol{r}} \otimes \boldsymbol{C}_{h,h}, \end{aligned} \tag{6.4}$$

$$\mathsf{E}\left[\bar{\boldsymbol{h}}_n\bar{\boldsymbol{h}}_n\right] = \underbrace{\begin{bmatrix} r_0 & r_1 & \cdots & r_{M-1} \\ r_{-1} & r_0 & \cdots & r_{M-2} \\ \vdots & \vdots & & \vdots \\ r_{1-M} & r_{2-M} & \vdots & r_0 \end{bmatrix}}_{\boldsymbol{R}} \otimes \boldsymbol{C}_{h,h}, \tag{6.5}$$

where $\boldsymbol{r}_n = J_0\left(2\pi\frac{f_{\mathsf{D}}}{f_{\mathsf{a}}}n\right)$ and $f_{\mathsf{a}} = \frac{B}{N}$.

Linear Predictor

Given the autocorrelation matrix $\boldsymbol{R} \in \mathbb{C}^{M\times M}$, the autocorrelation vector $\boldsymbol{r} \in \mathbb{C}^{M}$, and the pilot symbols $\boldsymbol{X}^{\mathsf{H}}\boldsymbol{X} = \sigma_{\mathsf{x}}^2 \boldsymbol{I}_N$, the linear predictor of the OFDM channel coefficients can be obtained as

$$\boxed{\hat{\boldsymbol{H}}_{\mathsf{LP}} = (\boldsymbol{r} \otimes \boldsymbol{C}_{H,H})^{\mathsf{H}} \left(\boldsymbol{R} \otimes \boldsymbol{C}_{H,H} + \tfrac{1}{\gamma}\boldsymbol{I}_{MN}\right)^{-1} \bar{\boldsymbol{H}}_{\mathsf{LS}},} \tag{6.6}$$

where

$$\bar{\boldsymbol{H}}_{\mathsf{LS}} = \left[\hat{\boldsymbol{H}}^{\mathsf{H}}_{\mathsf{LS},n}, \hat{\boldsymbol{H}}^{\mathsf{H}}_{\mathsf{LS},n-1}, \ldots, \hat{\boldsymbol{H}}^{\mathsf{H}}_{\mathsf{LS},n-M+1}\right]^{\mathsf{H}} \in \mathbb{C}^{MN}, \tag{6.7}$$

and

$$\begin{aligned} \bar{\boldsymbol{H}}^{(\mathsf{int})}_{\mathsf{LP}} \;=\; & \boldsymbol{F}_L(\boldsymbol{r} \otimes \boldsymbol{C}_{h,h})^{\mathsf{H}}(\boldsymbol{I}_M \otimes \boldsymbol{F}_{P\times L})^{\mathsf{H}} \times \\ & \times \left[(\boldsymbol{I}_M \otimes \boldsymbol{F}_{P\times L})(\boldsymbol{R} \otimes \boldsymbol{C}_{h,h})(\boldsymbol{I}_M \otimes \boldsymbol{F}_{P\times L})^{\mathsf{H}} + \frac{1}{\gamma}\boldsymbol{I}_{MP}\right]^{-1} \bar{\boldsymbol{H}}_{P,\mathsf{LS}}, \end{aligned} \tag{6.8}$$

respectively.[2]

[2] In non-fully loaded OFDM systems where $\boldsymbol{X} \rightarrow \boldsymbol{X}_P$.

End

Wrap Up

You have seen

... some basics of OFDM from a theoretical point of view
... some basics about channel estimation for time-invariant OFDM
... a very brief extension to the case of channel estimation for time-variant OFDM

You have not seen

... really implementable stuff
... sophisticated channel estimation approaches based on (assisted by) iterative (turbo) principles
... channel tracking concepts and other stuff of channel prediction

$\Rightarrow$ a list of references will be provided soon!