Chapter 2
Discrete-Time Flat-Fading System Model

In the present chapter, we introduce the discrete-time stationary Rayleigh flat-fading system model as it will be used throughout the following derivations. Hereby, we also discuss the limitations of the model and recall the definitions of operational and information theoretic capacity as a basis for our further derivations. In this chapter, we restrict to the single-input single-output (SISO) channel. In Chapter 8 the model is extended to the multiple-input multiple-output (MIMO) scenario, and in Chapter 9 the frequency-selective case is discussed.

Fig. 2.1 shows a basic block-diagram of a transmission system based on a flat-fading channel. This model is based on several simplifications as it only shows a discrete-time baseband representation as we will use in the following. This discrete-time representation corresponds to the symbol rate. Clearly, a symbol rate representation cannot be used to study timing- and frequency synchronization which would require a modeling with a sampling rate higher than the symbol rate to get a sufficient statistic. As timing- and frequency synchronization are out of scope of the present work, for the moment this model is of sufficient detail to show the main effects concerning capacity evaluations. In Section 2.3 we discuss the limitations of this model and in Section 9.1, we extend the model to the case of wide-sense stationary uncorrelated scattering (WSSUS) frequency-selective fading channels, starting from a continuous-time representation.

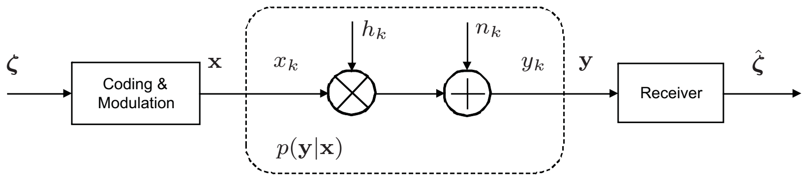

Fig. 2.1: Simplified block diagram of transmission system based on a discrete-time flat-fading channel

M. Dörpinghaus, *On the Achievable Rate of Stationary Fading Channels*, Foundations in signal Processing, Communications and Networking 6, DOI 10.1007/978-3-642-19780-2_2,

A binary data sequence $\boldsymbol{\zeta}$ of length N_{info} with the elements $\zeta_n \in \{0, 1\}$ is mapped by encoding and modulation to a transmit symbol sequence of length N represented by the vector $\mathbf{x}$ with the elements x_k. The transmit symbols x_k are corrupted by multiplication with the channel fading weight h_k and by additive noise n_k. Here, it is important to note that the channel fading weights h_k are assumed to be temporally correlated and the process $\{h_k\}$ is stationary. With respect to the discrete-time system model shown in Fig. 2.1, the channel output observations y_k are contained in the vector $\mathbf{y}$. From a mathematical point of view, the effect of the channel is completely described by the conditional probability density function (PDF) $p(\mathbf{y}|\mathbf{x})$. The aim of the receiver is to find an estimate $\hat{\boldsymbol{\zeta}}$ of the binary data sequence $\boldsymbol{\zeta}$ based on the observation vector $\mathbf{y}$. As we are interested in the evaluation of the channel capacity/achievable rate, we do not further discuss how the receiver acquires $\hat{\boldsymbol{\zeta}}$ based on the observation $\mathbf{y}$.

After this brief overview of the system model, we now focus on an exact mathematical representation of the system model required for further study. We consider an ergodic discrete-time jointly proper Gaussian [82] flat-fading channel, whose output at time k is given by

$$y_k = h_k x_k + n_k \tag{2.1}$$

where $x_k \in \mathbb{C}$ is the complex-valued channel input, $h_k \in \mathbb{C}$ represents the channel fading coefficient, and $n_k \in \mathbb{C}$ is additive white Gaussian noise. The processes $\{h_k\}$, $\{x_k\}$, and $\{n_k\}$ are assumed to be mutually independent.

We assume that the noise $\{n_k\}$ is a sequence of i.i.d. proper Gaussian random variables of zero-mean and variance σ_n^2. The stationary channel fading process $\{h_k\}$ is zero-mean jointly proper Gaussian. In addition, the fading process is time-selective and characterized by its autocorrelation function

$$r_h(l) = \mathrm{E}[h_{k+l} h_k^*]. \tag{2.2}$$

Its variance is given by $r_h(0) = \sigma_h^2$.

The normalized PSD of the channel fading process is defined by

$$S_h(f) = \sum_{l=-\infty}^{\infty} r_h(l) e^{-j2\pi l f}, \quad |f| < 0.5 \tag{2.3}$$

where we assume that the PSD exists and $j = \sqrt{-1}$. Here the frequency f is normalized with respect to the symbol duration T_{Sym}. In the following, we will in general use the normalized PSD and, thus, refer to it as PSD for simplification. For a jointly proper Gaussian process, the existence of the PSD implies ergodicity [103]. As the channel fading process $\{h_k\}$ is assumed to be stationary, $S_h(f)$ is real-valued. Because of the limitation of the velocity of the transmitter, the receiver, and of objects in the environment, the spread of the PSD is limited, and we assume it to be compactly supported within the interval $[-f_d, f_d]$, with $0 < f_d < 0.5$, i.e., $S_h(f) = 0$ for $f \notin [-f_d, f_d]$. The parameter f_d corresponds to the normalized maximum Doppler shift and, thus, indicates the dynamics of the channel. To ensure ergodicity, we exclude the case $f_d = 0$. Following the definition given in [24, Sec. XII.2, p. 564], this fading channel is sometimes referred to as *nonregular*.[1]

For technical reasons in some of the proofs, i.e., the calculation of the upper bound on the achievable data rate in Chapter 3, its extensions to the MIMO and the frequency selective case in Chapter 8 and Chapter 9, and the derivation of the lower bound on the achievable rate with joint processing of pilot and data symbols in Chapter 7, we restrict to autocorrelation functions $r_h(l)$ which are absolutely summable, i.e.,

$$\sum_{l=-\infty}^{\infty} |r_h(l)| < \infty \tag{2.4}$$

instead of the more general class of square summable autocorrelation functions, i.e.,

$$\sum_{l=-\infty}^{\infty} |r_h(l)|^2 < \infty \tag{2.5}$$

which is already fulfilled due to our assumption that the PSD exists. However, the assumption of absolutely summable autocorrelation functions is not a severe restriction. E.g., the important rectangular PSD, see (2.8), can be arbitrarily closely approximated by a PSD with the shape corresponding to the transfer function of a raised cosine filter, whose corresponding autocorrelation function is absolutely summable.

2.1 Rayleigh Fading and Jakes' Model

The assumed Rayleigh fading model is a commonly used fading model which reasonably describes the channel observed in mobile urban environments with many scattering objects and no line of sight, see, e.g., [15]. It has to be mentioned that the Rayleigh model only describes small scale fading. Large scale fading due to path loss and shadowing is not described by this model and is also outside the scope of this thesis.

Due to the assumption of relative motion with constant velocity between transmitter, receiver, and objects in the environment, the fading becomes temporally correlated. The normalized continuous-time autocorrelation function is given by [54], [112]

[1] For a discussion on the justification if a physical channel fading process is *nonregular* based on *real world numbers* see [66].

$$r(t) = \sigma_h^2 J_0 \left(2\pi f_d \frac{t}{T_{\text{Sym}}} \right) \tag{2.6}$$

where $J_0(\cdot)$ is a zeroth-order Bessel function of the first kind. The corresponding PSD $S_h(f)$ in (2.3) of the discrete-time fading process is given by

$$S_h(f)\big|_{\text{Jakes}} = \begin{cases} \frac{\sigma_h^2}{\pi\sqrt{f_d^2 - f^2}} & \text{for } |f| < f_d \\ 0 & \text{for } f_d \leq |f| \leq 0.5 \end{cases} . \tag{2.7}$$

These correlation properties can be derived analytically for a dense scatterer environment with a vertical receive antenna with a constant azimuthal gain, a uniform distribution of signals arriving at all angles, i.e., in the interval $[0, 2\pi)$, and with uniformly distributed phases based on a sum of sinusoids [54]. The sum-of-sinusoids model can also be used to generate temporally correlated Rayleigh fading for simulation.

Sometimes the Jakes' PSD in (2.7) is approximated by the following rectangular PSD

$$S_h(f)\big|_{\text{rect}} = \begin{cases} \frac{\sigma_h^2}{2f_d} & \text{for } |f| \leq f_d \\ 0 & \text{for } f_d < |f| \leq 0.5 \end{cases} . \tag{2.8}$$

This approximation entails only a small difference with respect to the performance of the corresponding channel estimation [6], [79, pp. 651 and 658]. The performance of the channel estimation can be measured by the estimation error variance, which is related to the capacity. For mathematical tractability, we will also use the rectangular PSD in (2.8) for several derivations.

As already stated, the discrete-time autocorrelation function $r_h(l)$ corresponding to a rectangular PSD $S_h(f)\big|_{\text{rect}}$, which is given by

$$r_h(l) = \sigma_h^2 \text{sinc}(2 f_d l) \tag{2.9}$$

is not absolutely summable. However, the rectangular PSD can be arbitrarily closely approximated by a PSD with a shape corresponding to the transfer function of a raised cosine (RC) filter, i.e.,

$$S_h(f)\big|_{\text{RC}} = \begin{cases} \frac{\sigma_h^2}{2f_d} & \text{for } |f| \leq (1-\beta_{\text{ro}}) f_d \\ \frac{\sigma_h^2}{4f_d}\left[1 - \sin\left(\frac{\frac{2\pi}{2f_d}(f-f_d)}{2\beta_{\text{ro}}}\right)\right] & \text{for } (1-\beta_{\text{ro}}) f_d < f \leq (1+\beta_{\text{ro}}) f_d \\ 0 & \text{for } f_d(1+\beta_{\text{ro}}) < |f| \leq 0.5 \end{cases} . \tag{2.10}$$

Here $0 \leq \beta_{\text{ro}} \leq 1$ is the roll-off factor. For $\beta_{\text{ro}} \to 0$ the PSD $S_h(f)\big|_{\text{RC}}$ approaches the rectangular PSD $S_h(f)\big|_{\text{rect}}$. Furthermore, the discrete-time autocorrelation function corresponding to $S_h(f)\big|_{\text{RC}}$ is given by

$$r_h(l) = \sigma_h^2 \mathrm{sinc}(2f_d l) \frac{\cos(\beta_{\mathrm{ro}} \pi 2 f_d l)}{1 - 4\beta_{\mathrm{ro}}^2 4 f_d^2 l^2} \tag{2.11}$$

which for $\beta_{\mathrm{ro}} > 0$ is absolutely summable. Thus, the rectangular PSD in (2.8) can be arbitrarily closely approximated by a PSD with an absolutely summable autocorrelation function. Therefore, in the rest of this work, we often evaluate the derived bounds on the achievable rate for a rectangular PSD of the channel fading process, although some of the derivations are based on the assumption of an absolutely summable autocorrelation function.

Typical fading channels, as they are observed in mobile communication environments, are characterized by relatively small normalized Doppler frequencies f_d in the regime of $f_d \ll 0.1$. Therefore, the restriction to channels with $f_d < 0.5$, i.e., *nonregular* fading, in the present work is reasonable. Depending on the carrier frequency f_c, the relative velocity between transmitter and receiver v, and the symbol rate $1/T_{\mathrm{Sym}}$, the maximum normalized Doppler frequency is given by

$$f_d = \frac{v \cdot f_c}{c} \cdot T_{\mathrm{Sym}} \tag{2.12}$$

where $c \approx 2.998 \cdot 10^8$ m/s is the speed of light. Considering parameters of a typical mobile communication system, e.g., $f_c = 2$ GHz, a maximum velocity of $v = 300$ km/h, and a symbol rate of 1 MHz leads to a maximum Doppler frequency of only $f_d \approx 0.00056$.

Notice that although in practical scenarios the observed channel dynamics are very small, within this work we always consider the range of $0 < f_d < 0.5$ to get a thorough understanding of the behavior of the bounds on the achievable rate.

2.2 Matrix-Vector Notation

We base the derivation of bounds on the achievable rate on the following matrix-vector notation of the system model:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} = \mathbf{X}\mathbf{h} + \mathbf{n} \tag{2.13}$$

where the vectors are defined as

$$\mathbf{y} = [y_1, \ldots, y_N]^T \tag{2.14}$$

$$\mathbf{x} = [x_1, \ldots, x_N]^T \tag{2.15}$$

$$\mathbf{n} = [n_1, \ldots, n_N]^T. \tag{2.16}$$

The matrix $\mathbf{H}$ is diagonal and defined as $\mathbf{H} = \mathrm{diag}(\mathbf{h})$ with $\mathbf{h} = [h_1, \ldots, h_N]^T$. Here the diag$(\cdot)$ operator generates a diagonal matrix whose diagonal elements are given by the argument vector. The diagonal matrix $\mathbf{X}$ is given by $\mathbf{X} = \mathrm{diag}(\mathbf{x})$. The quantity N is the number of considered symbols. Later on, we investigate the case of $N \to \infty$ to evaluate the achievable rate.

Using this vector notation, we express the temporal correlation of the fading process by the correlation matrix

$$\mathbf{R}_h = \mathrm{E}[\mathbf{h}\mathbf{h}^H] \tag{2.17}$$

which has a Hermitian Toeplitz structure.

Concerning the input distribution, unless otherwise stated, we make the assumption that the symbols x_k are i.d. , with an maximum average power σ_x^2. For the nominal mean SNR we introduce the variable

$$\rho = \frac{\sigma_x^2 \sigma_h^2}{\sigma_n^2}. \tag{2.18}$$

Notice that we use the term *nominal* mean SNR as ρ only corresponds to the actual mean SNR in case σ_x^2 is the average transmit power. For the case of a non-peak power constrained input distribution, the achievable rate is maximized by using the maximum average transmit power σ_x^2. Thus, in the non-peak power constrained case ρ corresponds to the actual mean SNR.

2.3 Limitations of the Model

In this section, we discuss the limitations of the model. Therefore, let us consider an appropriately bandlimited continuous-time model first, where the channel output is given by

$$y(t) = h(t)s(t) + n(t) \tag{2.19}$$

with $h(t)$ being the continuous-time channel fading process, i.e., the corresponding discrete-time process h_k is given by

$$h_k = h(kT_{\mathrm{Sym}}) \tag{2.20}$$

where T_{Sym} is the symbol duration. Analogously, the continuous-time and the discrete-time additive noise and channel output processes are related by

$$n_k = n(kT_{\mathrm{Sym}}) \tag{2.21}$$

$$y_k = y(kT_{\mathrm{Sym}}). \tag{2.22}$$

The continuous-time transmit process $s(t)$ is given by

$$s(t) = \sum_{k=-\infty}^{\infty} x_k g(t - kT_{\mathrm{Sym}}) \tag{2.23}$$

where $g(t)$ is the transmit pulse. We assume the use of bandlimited transmit pulses, which have an infinite impulse response[2]. In typical systems often root-raised cosine pulses are used such that in combination with the matched filter at the receiver intersymbol interference is minimized. Their impulse response $g(t)$ and their normalized frequency response $G(f)$ are given by

$$G(f) = \sqrt{G_{\mathrm{RC}}(f)} \tag{2.24}$$

with $G_{\mathrm{RC}}(f)$ being the transfer function of the raised cosine filter, cf. (2.10)

$$G_{\mathrm{RC}}(f) = \begin{cases} T_{\mathrm{Sym}} & \text{for } |f| \leq \frac{1-\beta_{\mathrm{ro}}}{2} \\ \frac{T_{\mathrm{Sym}}}{2}\left[1 + \cos\left(\frac{\pi}{\beta_{\mathrm{ro}}}\left[|f| - \frac{1-\beta_{\mathrm{ro}}}{2}\right]\right)\right] & \text{for } \frac{1-\beta_{\mathrm{ro}}}{2} < |f| \leq \frac{1+\beta_{\mathrm{ro}}}{2} \\ 0 & \text{otherwise} \end{cases} \tag{2.25}$$

and [3]

$$g(t) = \begin{cases} 1 - \beta_{\mathrm{ro}} + 4\frac{\beta_{\mathrm{ro}}}{\pi} & \text{for } t = 0 \\ \frac{\beta_{\mathrm{ro}}}{\sqrt{2}}\left[\left(1 + \frac{2}{\pi}\right)\sin\left(\frac{\pi}{4\beta_{\mathrm{ro}}}\right) + \left(1 - \frac{2}{\pi}\right)\cos\left(\frac{\pi}{4\beta_{\mathrm{ro}}}\right)\right] & \text{for } t = \pm\frac{T_{\mathrm{Sym}}}{4\beta_{\mathrm{ro}}} \\ \frac{\sin\left[\pi\frac{t}{T_{\mathrm{Sym}}}(1-\beta_{\mathrm{ro}})\right] + 4\beta_{\mathrm{ro}}\frac{t}{T_{\mathrm{Sym}}}\cos\left[\pi\frac{t}{T_{\mathrm{Sym}}}(1+\beta_{\mathrm{ro}})\right]}{\pi\frac{t}{T_{\mathrm{Sym}}}\left[1-\left(4\beta_{\mathrm{ro}}\frac{t}{T_{\mathrm{Sym}}}\right)^2\right]} & \text{otherwise} \end{cases} . \tag{2.26}$$

In the following, we assume a roll-off factor of $\beta_{\mathrm{ro}} = 0$ corresponding to sinc transmit pulses and, thus, we have a rectangular PSD with a normalized bandwidth of 1.

The continuous-time input/output relation in (2.19) has the following stochastic representation in frequency domain

$$S_y(f) = S_h(f) \star S_s(f) + S_n(f) \tag{2.27}$$

where $\star$ denotes convolution and $S_y(f)$, $S_h(f)$, $S_s(f)$, and $S_n(f)$ are the normalized power spectral densities of the continuous-time processes $y(t)$, $h(t)$, $s(t)$, and $n(t)$, e.g.,

$$S_s(f) = \int_{-\infty}^{\infty} \mathrm{E}\left[s(t+\tau)s^*(t)\right] e^{-j2\pi f\tau} d\tau \tag{2.28}$$

[2] For a discussion on the contradiction that physical signals must be bandlimited but on the other hand are not infinite in time the interested reader is referred to [113].

and correspondingly for the other PSDs. Here we always assume normalization with $1/T_{\mathrm{Sym}}$.

We are interested in the normalized bandwidth of the component $S_h(f) \star S_s(f)$, i.e., the component containing information on the transmitted sequence $\{x_k\}$. The normalized bandwidth of the transmit signal $s(t)$ directly corresponds to the normalized bandwidth of the transmit pulse $g(t)$, which is assumed to be 1, see above. The normalized bandwidth of the channel fading process is given by $2f_d$. Thus, the normalized bandwidth of the component $S_h(f) \star S_s(f)$ is given by $1 + 2f_d$.

To get a sufficient statistic, we would have to sample the channel output $y(t)$ at least with a frequency of $\frac{1+2f_d}{T_{Sym}}$. As the discrete-time channel output process $\{y_k\}$ is a sampled version of $y(t)$ with the rate $1/T_{\mathrm{Sym}}$, the discrete-time observation process $\{y_k\}$ is not a sufficient statistic of $y(t)$. As usually the normalized maximum Doppler frequency f_d is very small, the amount of discarded information is negligible.

All further derivations are based on the discrete-time model and therefore, are not based on a sufficient statistic, i.e., information is discarded. Beside the fact that in realistic systems the dynamics is very small and, thus, the amount of discarded information is small, in typical systems channel estimation is also performed at symbol rate signals and therefore also exhibits the loss due to the lack of a sufficient statistic. In addition, much of the current literature on the study of the capacity of stationary Rayleigh fading channels, e.g., [67], and on the achievable rate with synchronized detection, e.g., [6], is based on symbol rate discrete-time input-output relations and therefore do not ask the question about a sufficient statistic. However, this should not be understood as a motivation to use the symbol rate signal model. Furthermore, these considerations should be kept in mind in the later evaluations, especially, as we examine the derived bounds not only for very small values of f_d.

2.4 Operational and Information Theoretic Capacity

In this section, we first briefly introduce the definition of information theoretic capacity given by Shannon and recall the channel coding theorem [108] linking information theoretic capacity to operational capacity. Furthermore, we recall results on the extension of Shannon's definition of capacity, which was restricted to memoryless channels, to channels with memory, like the stationary fading channel considered in the present work.

For a memoryless channel the *information theoretic capacity* C_{info} is defined as the maximum mutual information[3] $\mathcal{I}(y; x)$ between the channel input x and the channel output y while maximizing over the distribution of the in-

[3] Notice that Shannon [108] defined the capacity without using the mutual information, but directly based on the definition of entropy.

put sample x, i.e., see also [19]

$$C_{\text{info}} = \max_{p(x)} \mathcal{I}(y;x) \tag{2.29}$$

where the mutual information is given by

$$\mathcal{I}(y;x) = h(y) - h(y|x) \tag{2.30}$$

and $h(\cdot)$ is the differential entropy defined by

$$h(y) = \mathrm{E}_y \left[\log\left(p(y)\right)\right] \tag{2.31}$$

$$h(y|x) = \mathrm{E}_{y,x} \left[\log\left(p(y|x)\right)\right]. \tag{2.32}$$

Based on (2.30) the mutual information $\mathcal{I}(y;x)$ can be understood as a measure about the information the random variable y contains about the random variable x, or alternatively on the reduction of the uncertainty on y when knowing x.

Besides this mathematical definition of capacity the main contribution of Shannon was the channel coding theorem. With the channel coding theorem Shannon proved that for memoryless channels all rates below the capacity C_{op}—which in this context often is named *operational capacity*— are achievable, i.e., for each rate $R < C_{\text{op}}$ there exists a code for which the probability of an erroneously decoded codeword approaches zero in the limit of an infinite codeword length. Conversely, this means that error-free transmission is only possible at rates R with $R \leq C_{\text{op}}$. The channel coding theorem states that for memoryless channels, information theoretic and operational capacity coincide. Therefore, the capacity C for a memoryless channel is given by

$$C \equiv C_{\text{info}} = C_{\text{op}}. \tag{2.33}$$

Depending on the specific type of channel, it is necessary to introduce further constraints, to get a finite capacity. E.g., for the additive white Gaussian noise (AWGN) channel the capacity is infinite in the case of an unconstrained input power. Therefore, usually its capacity is given based on a constraint on the maximum average input power P_{av}.

In the context of the stationary flat-fading channel model given in (2.1) and (2.13), we introduce the following constraints on the average power P_{av} and the peak power P_{peak},

$$\frac{1}{N}\mathrm{E}\left[\mathbf{x}^H\mathbf{x}\right] \leq P_{\text{av}} \tag{2.34}$$

$$\max_{1\leq k\leq N} |x_k|^2 \leq P_{\text{peak}}. \tag{2.35}$$

The average power constraint will be always used in the following derivations, while we will use the peak power constraint only at specific places.

It is important to note that Shannon's theorem is based on the assumption of memoryless channels, i.e., all usages of the channel are independent. For the stationary fading channel considered in the present work, this assumption does not hold, as the channel fading process is temporally correlated. Due to this temporal correlation, e.g., the channel observation y_k also contains information on the channel fading weight h_{k-1} and thus on the previous transmit symbol x_{k-1}.

The coincidence of information theoretic capacity and operational capacity can be extended to channels with memory under some further conditions [103]. Before discussing this, we introduce the definition of information theoretic capacity in the context of the stationary fading channel given in (2.13).

The information theoretic capacity per unit time of the stationary fading channel model is given by

$$C_{\text{info}} = \lim_{N\to\infty} \sup_{\mathcal{P}_{\text{gen}}} \frac{1}{N} \mathcal{I}(\mathbf{y};\mathbf{x}). \tag{2.36}$$

where the supremum is taken over the set $\mathcal{P}_{\text{gen}}$ of input distributions given by

$$\mathcal{P}_{\text{gen}} = \left\{ p(\mathbf{x}) \middle| \mathbf{x} \in \mathbb{C}^N, \frac{1}{N}\mathrm{E}[\mathbf{x}^H\mathbf{x}] \leq P_{\text{av}}, \max_{1\leq k\leq N} |x_k|^2 \leq P_{\text{peak}} \right\}. \tag{2.37}$$

The definition of the information theoretic capacity holds whenever the limit in (2.36) exists.

The peak-power constraint in (2.37) is not generally necessary. Only some of the following derivations are based on a peak power constraint. The case of an unconstrained peak-power corresponds to $P_{\text{peak}} = \infty$ in (2.37).

Corresponding to the memoryless channel, in case of the stationary fading channel the operational capacity C_{op} corresponds to the maximum achievable rate R, which implies the existence of a code with a decoding error probability that approaches zero for infinite codeword length, i.e., $N \to \infty$.

Now we recall the conditions required for the coincidence of information theoretic and operational capacity in case of a channel with memory given in [103]. To describe these conditions, we quote the following definitions on weakly mixing and ergodic processes given in [103], which itself cites [74, Section 5] and [91, p.70].

Define $\phi_i(z_1, z_2, \ldots, z_n)$ with $i = 1, 2$ to be two bounded measurable functions of an arbitrary number of complex variables $z_1, \ldots, z_n$. Furthermore, define the operator M_t as $\lim_{t\to\infty} \frac{1}{t}\sum_1^t$ for a discrete-time process $\{z_k\}$. In addition we define $\psi(t)$ as

$$\begin{aligned}\psi(t) &= \mathrm{E}\left[\phi_1(z_{t_1}, \ldots, z_{t_n}) \cdot \phi_2(z_{t_1^*+t}, \ldots, z_{t_n^*+t})\right] \\ &\quad - \mathrm{E}\left[\phi_1(z_{t_1}, \ldots, z_{t_n}) \cdot \phi_2(z_{t_1^*}, \ldots, z_{t_n^*})\right].\end{aligned} \tag{2.38}$$

A stationary stochastic process $\{z_k\}$ is

- weakly mixing if, for all choices of ϕ_1, ϕ_2 and times $t_1, \dots, t_n, t_1^*, \dots, t_n^*$

$$M_t\left[\psi^2(t)\right] = 0 \tag{2.39}$$

- ergodic if, for all choices of ϕ_1, ϕ_2 and times $t_1, \dots, t_n, t_1^*, \dots, t_n^*$

$$M_t\left[\psi(t)\right] = 0. \tag{2.40}$$

Notice, that an ergodic process is also weakly mixing. Based on concepts concerning information stability and the Shannon-McMillan-Breiman theorem for finite-alphabet ergodic sources, the following proposition is derived in [103]:

Proposition [103]: If the processes $\{h_k\}$ and $\{n_k\}$ are stationary weakly mixing, and if $\{h_k\}$, $\{n_k\}$, and $\{x_k\}$ are mutually independent, then for every $P_{\text{av}}, P_{\text{peak}} > 0$ the information theoretic capacity C_{info} is well defined and corresponds to the operational capacity C_{op}.

As we assume that the PSD of the channel fading process $S_h(f)$ in (2.3) exists, and as the fading process is assumed to be jointly proper Gaussian, the channel fading process is ergodic. For a discussion on this relation see [103]. For proper Gaussian processes, ergodicity is equivalent to weakly mixing. Thus, for the system model (2.13) considered in this work, operational and information theoretic capacity coincide. This allows us to use the term of *information theoretic capacity* in the following.

2.4.1 Outage Capacity

For completeness of presentation, we also mention that there exist further capacity measures. The preceding definition of information theoretic capacity (*Shannon capacity*) considers the maximum rate being achievable with a probability of an decoding error that approaches zero for infinitely long codewords. If we deviate from the focus on an arbitrary small probability of error, we can also use the definition of outage capacity.

The q%-outage capacity C_{outage} is defined as the information rate that is guaranteed for $(100 - q)$% of the channel realizations [84], i.e.,

$$P(C \leq C_{\text{outage}}) = q\%. \tag{2.41}$$

Therefore, this definition is especially interesting in the context of channels, where the channel quality changes over time like fading channels.

However, within the rest of this work we will not use the measure *outage capacity* but will restrict to the use of the *information theoretic capacity.*

On the Achievable Rate of Stationary Fading Channels
Dörpinghaus, M.
2011, XIV, 310 p., Hardcover
ISBN: 978-3-642-19779-6