Practical Wireless Communications (WORK IN PROGRESS)

Beamforming, Receiver Dynamic Range, OFDM, SINR, Link Adaptation, and Error Recovery

Roman Grigorii

Ph.D., Robotics and Human–Machine Interaction
Senior Software Engineer, Robotics

2026

Preface

During my time at Tarana Wireless, I had the opportunity to work on fixed-wireless communication systems and learn a great deal about how modern radios actually work in practice. My work touched areas such as beamforming, link acquisition, receiver gain control, interference, and physical-layer performance.

Many of these topics are easy to encounter as isolated engineering problems without ever stepping back to connect them to the underlying communications theory. This document is my attempt to do that. I wanted to organize what I learned, preserve some of the practical intuition that came from working on real systems, and build a reference that I can return to in the future.

Although much of the material comes from problems I encountered in wireless communications, many of the same ideas: estimation, feedback, signal processing, dynamic range, and adaptation also carry over naturally into robotics and embedded systems.

I wrote this primarily as a reference for myself, but I hope it is useful to others as well.

Contents

Chapter 1

Fixed Wireless Communication Systems

1.1 Internet Service Providers and the Last Mile

An Internet Service Provider, or ISP, is an organization that provides individuals, businesses, and institutions with connectivity to the global Internet. From a high-level perspective, an ISP operates infrastructure that connects local users to larger regional and global networks. This infrastructure may include fiber-optic backbone links, data centers, switching and routing equipment, wireless base stations, and interconnections with other network operators.

The Internet itself is not a single centrally managed network. It is a collection of many independently operated networks that exchange traffic with one another. ISPs form an important part of this structure by providing the local and regional connectivity that allows end users to reach the broader Internet.

A simplified view of Internet connectivity is

$$\text{user device} \rightarrow \text{local network} \rightarrow \text{ISP access network} \rightarrow \text{ISP backbone} \rightarrow \text{Internet}.$$

The connection between an ISP's high-capacity infrastructure and the individual customer is commonly referred to as the *last mile*. Despite the name, the last mile may physically extend for several meters, several kilometers, or even considerably farther depending on geography and network architecture.

The last mile is often one of the most difficult and expensive parts of a communication network to deploy. High-capacity backbone links can aggregate traffic from thousands or millions of users, while the last-mile infrastructure must ultimately reach each individual home, business, or device.

For example, constructing a fiber backbone between two cities may provide enormous aggregate capacity. However, connecting every individual building in those cities requires additional fiber runs, trenches, utility access, installation labor, and customer-premises equipment. The cost is therefore distributed across a much smaller number of users at the edge of the network.

Several technologies can be used to implement the last mile, including

- fiber-optic communication,
- coaxial cable,
- Digital Subscriber Line (DSL),

- cellular radio,
- satellite communication,
- fixed wireless access.

Each technology represents a different compromise among bandwidth, deployment cost, range, latency, reliability, and infrastructure requirements.

1.1.1 Fiber-Optic Access

Fiber-optic communication carries information using light transmitted through glass or plastic fibers. Fiber provides extremely high bandwidth and low signal attenuation, making it one of the most capable technologies for broadband access.

A simplified fiber connection can be represented as

$$\text{ISP fiber network} \rightarrow \text{customer optical terminal} \rightarrow \text{local Ethernet/Wi-Fi network}.$$

The principal disadvantage of fiber is deployment cost. New fiber may require trenching, utility-pole access, permits, splicing, and installation at each customer location. In densely populated areas these costs can be shared across many subscribers, but in rural or difficult terrain the cost per customer may become substantial.

1.1.2 Cable and DSL

Cable networks commonly use coaxial cable infrastructure originally deployed for television distribution. DSL systems use existing copper telephone wiring.

These technologies have the advantage of reusing physical infrastructure that may already reach individual buildings. However, their performance is ultimately limited by the characteristics of the underlying transmission medium.

Copper-based systems generally experience greater attenuation and bandwidth limitations than modern fiber systems, especially over long distances.

1.1.3 Cellular Access

Cellular systems provide Internet connectivity through mobile radio networks. A user device communicates with a nearby cell site, which then connects to the operator's core network.

The architecture can be approximated as

$$\text{mobile device} \leftrightarrow \text{cellular base station} \rightarrow \text{operator core network} \rightarrow \text{Internet}.$$

Cellular systems are designed primarily to support mobile users whose position may change continuously.

This requirement creates different design constraints from fixed wireless systems, where the transmitter and receiver are installed at known locations and remain approximately stationary.

1.1.4 Satellite Access

Satellite systems transmit signals between a ground terminal and a satellite in orbit.

A simplified connection is

$$\text{customer terminal} \leftrightarrow \text{satellite} \leftrightarrow \text{ground infrastructure} \rightarrow \text{Internet}.$$

Satellite systems can provide service in locations where terrestrial infrastructure is difficult to deploy. However, their performance is influenced by factors such as orbital altitude, propagation delay, atmospheric conditions, antenna pointing, and available spectrum.

1.2 Fixed Wireless Access

Fixed Wireless Access, commonly abbreviated as FWA, provides broadband connectivity through a radio link between fixed endpoints.

Instead of physically extending a cable to the customer, an ISP installs a radio at the customer location and communicates with it from a base station or tower.

A simplified architecture is

$$\text{Internet backbone} \rightarrow \text{base station} \rightarrow \text{wireless link} \rightarrow \text{customer radio} \rightarrow \text{local network}.$$

This architecture can significantly reduce deployment cost because the communication medium between the base station and customer is free space rather than a newly installed cable.

The physical installation may therefore require only

- a base station connected to the ISP network,
- an antenna or radio installed at the customer location,
- power,
- configuration and alignment.

Once the base station infrastructure exists, additional customers can often be added without constructing an entirely new physical path back to the ISP.

This economic advantage is one of the principal motivations for fixed wireless access.

1.3 Base Nodes and Remote Nodes

In systems such as those developed by Tarana Wireless, the network-side radio is commonly referred to as a *Base Node*, or BN, while the customer-side radio is referred to as a *Remote Node*, or RN.

The fundamental communication link can therefore be represented as

$$\boxed{\text{BN} \longleftrightarrow \text{wireless channel} \longleftrightarrow \text{RN}}$$

The Base Node is typically connected to the ISP's higher-capacity network infrastructure. It may serve multiple customers located within a geographic sector.

The Remote Node is installed at or near the customer premises and provides the wireless endpoint for that customer's broadband connection.

Traffic therefore flows in both directions:

$$\text{BN} \rightarrow \text{RN}$$

for downlink traffic, and

$$\text{RN} \rightarrow \text{BN}$$

for uplink traffic.

The base station must generally support communication with many remote nodes, each of which may experience a different radio channel.

Thus, even if the BN hardware is common to all customers, the propagation environment associated with each RN may be significantly different.

One remote node may have a nearly unobstructed path to the base station, while another may communicate primarily through reflections from buildings or other structures.

This is one reason why sophisticated physical-layer processing is particularly valuable in modern fixed wireless systems.

1.4 Why Fixed Wireless Is Difficult

At first glance, fixed wireless communication might appear easier than mobile communication. Since both the base station and customer radio remain physically stationary, one might expect the radio channel to remain essentially constant.

In practice, this is not the case.

A useful principle is

$$\boxed{\text{fixed endpoints do not imply a fixed wireless channel.}}$$

The electromagnetic environment surrounding the radios can change even when the radios themselves do not move.

The transmitted signal interacts with buildings, vegetation, vehicles, terrain, and other objects in the environment before reaching the receiver.

As a result, the receiver often observes not one copy of the transmitted signal, but many delayed and attenuated copies.

1.5 Multipath Propagation

Consider a transmitter and receiver with a direct line-of-sight path:

$$\text{TX} \rightarrow \text{RX}.$$

This direct path is only one possible route by which electromagnetic energy may reach the receiver.

A signal may reflect from a building:

$$\text{TX} \rightarrow \text{building} \rightarrow \text{RX}.$$

It may also reflect from the ground:

$$\text{TX} \rightarrow \text{ground} \rightarrow \text{RX}.$$

Additional paths may involve multiple reflections:

$$\text{TX} \to \text{building} \to \text{ground} \to \text{RX}.$$

Each path generally has a different propagation distance.
Since electromagnetic waves travel at approximately the speed of light,

$$c \approx 3 \times 10^8 \text{ m/s},$$

the propagation delay associated with a path of length d is approximately

$$\tau = \frac{d}{c}.$$

If two paths have different lengths, they arrive at the receiver at different times.
Suppose one path has propagation delay

$$\tau_0$$

and another has delay

$$\tau_1.$$

The receiver therefore observes multiple copies of the same transmitted waveform:

$$y(t) = a_0 x(t - \tau_0) + a_1 x(t - \tau_1) + \cdots,$$

where a_0 and a_1 represent the attenuation and phase modification associated with each path.
More generally, the received signal can be written as

$$y(t) = \sum_{l=0}^{L-1} h_l x(t - \tau_l) + n(t),$$

where

- L is the number of significant propagation paths,
- h_l is the complex gain associated with path l,
- τ_l is its delay,
- $n(t)$ represents receiver noise.

This phenomenon is known as *multipath propagation*.

1.6 Constructive and Destructive Interference

Because radio signals are waves, the multiple received copies may either reinforce or cancel one another.

Suppose two received waves have approximately equal amplitude.
If they arrive with similar phase,

$$x_1(t) + x_2(t)$$

may produce a larger total amplitude.

This is constructive interference.

If they arrive approximately 180° out of phase,

$$x_1(t) + x_2(t) \approx 0,$$

the waves may partially cancel.

This is destructive interference.

As a result, very small changes in path length can produce significant changes in received signal strength.

For a carrier with wavelength

$$\lambda = \frac{c}{f},$$

a path-length change of

$$\frac{\lambda}{2}$$

corresponds to approximately 180° of phase change.

At gigahertz frequencies, the wavelength may be only a few centimeters.

Consequently, relatively small environmental changes can significantly change the effective channel.

1.7 Frequency-Selective Fading

Multipath does not necessarily affect every frequency in the same way.

Different propagation paths introduce different delays, and therefore different phase shifts as a function of frequency.

The channel can therefore be represented as a frequency-dependent quantity

$$H(f).$$

At one frequency,

$$|H(f_1)|$$

may be large, while at another nearby frequency,

$$|H(f_2)|$$

may be much smaller.

The channel is then said to exhibit *frequency-selective fading*.

This phenomenon is one of the major motivations for techniques such as OFDM, which divide a wideband channel into many narrowband subchannels.

1.8 Foliage and Environmental Obstructions

Trees and vegetation can strongly influence wireless propagation.

Leaves, branches, and moisture can

- absorb electromagnetic energy,
- scatter the signal,
- introduce additional multipath,
- change with wind and weather.

Even if the transmitter and receiver remain fixed, movement of nearby foliage can alter the propagation channel over time.

The channel therefore becomes time varying:

$$H(f, t).$$

The rate of variation may be much slower than in a mobile cellular system, but it can still be significant for beamforming, channel estimation, and link adaptation.

1.9 Moving Objects

Vehicles, machinery, people, and other moving objects can create additional reflected paths.

For example,

$$\text{TX} \to \text{vehicle} \to \text{RX}$$

may temporarily create a strong reflection.

As the vehicle moves, the path length changes, which alters both the delay and phase of the reflected signal.

The corresponding channel coefficient may therefore evolve with time.

1.10 Atmospheric Effects

Radio propagation can also be influenced by atmospheric conditions.

Depending on frequency and link length, relevant effects may include

- rain attenuation,
- humidity,
- atmospheric absorption,
- temperature-dependent propagation changes.

The severity of these effects depends strongly on carrier frequency.

For many terrestrial broadband systems, multipath and interference may dominate under normal conditions, but atmospheric effects remain part of the link budget and reliability analysis.

1.11 Interference

A receiver rarely operates in isolation.

Other transmitters may occupy the same or nearby spectrum.

A more realistic received-signal model is therefore

$$y(t) = h_s(t) * x_s(t) + \sum_{k=1}^{K} h_k(t) * x_k(t) + n(t),$$

where

- $x_s(t)$ is the desired signal,
- $h_s(t)$ is the desired channel,
- $x_k(t)$ represents interfering transmitters,
- $h_k(t)$ represents their corresponding channels,
- $n(t)$ is receiver noise.

The receiver must therefore distinguish the desired transmission not merely from noise, but also from structured signals transmitted by other radios.

This distinction motivates the use of Signal-to-Interference-plus-Noise Ratio rather than signal strength alone.

The SINR is

$$\text{SINR} = \frac{P_S}{P_I + P_N},$$

where

- P_S is desired-signal power,
- P_I is interference power,
- P_N is noise power.

A strong received signal does not necessarily imply a good communication link if interference is also strong.

1.12 Oscillator Mismatch and Frequency Error

The transmitter and receiver each use local oscillators to generate reference frequencies.

No physical oscillator is perfectly accurate.

Suppose the transmitter carrier is

$$f_{\text{TX}}$$

and the receiver assumes

$$f_{\text{RX}}.$$

The carrier-frequency offset is

$$\Delta f = f_{\mathrm{TX}} - f_{\mathrm{RX}}.$$

Even small frequency errors can accumulate significant phase rotation over time.

The received phase error evolves approximately as

$$\phi(t) = 2\pi \Delta f \, t.$$

This can interfere with coherent demodulation and, in multicarrier systems such as OFDM, can destroy the orthogonality between neighboring subcarriers.

The receiver therefore requires synchronization algorithms that estimate and compensate for frequency and phase errors.

1.13 RF Hardware Imperfections

The mathematical communication model often assumes ideal transmitters and receivers.

Real hardware introduces additional impairments.

Examples include

- amplifier nonlinearity,
- ADC quantization,
- clipping,
- phase noise,
- IQ imbalance,
- gain mismatch,
- timing error,
- antenna mismatch.

These effects become particularly important in high-order modulation schemes, where constellation points are closely spaced.

For example, a system using 256-QAM requires substantially more accurate amplitude and phase recovery than a system using QPSK.

The practical receiver must therefore compensate not only for the wireless channel but also for its own imperfections.

1.14 Why Digital Signal Processing Is Essential

The purpose of the modern wireless receiver is ultimately to recover transmitted information from a signal that has been distorted by propagation, interference, noise, and hardware imperfections.

A simplified receiver chain may contain

RF signal $\rightarrow$ ADC $\rightarrow$ synchronization $\rightarrow$ channel estimation $\rightarrow$ equalization $\rightarrow$ demodulation $\rightarrow$ error correction.

In multi-antenna systems, beamforming may occur before or during these operations:

$$\text{antenna array} \rightarrow \text{beamforming} \rightarrow \text{channel processing}.$$

Several major functions are therefore required.

1.14.1 Synchronization

The receiver must determine when the transmitted waveform begins and align its processing with the transmitter.

It may need to estimate

- symbol timing,
- frame timing,
- carrier-frequency offset,
- phase offset.

1.14.2 Channel Estimation

The receiver must estimate how the propagation environment has modified the transmitted signal.

In a simple narrowband model,

$$y = hx + n,$$

the receiver estimates

$$\hat{h}.$$

In an OFDM system, it may estimate

$$\hat{H}[k]$$

for each relevant subcarrier.

1.14.3 Equalization

Once the channel is estimated, the receiver attempts to remove its effect.

For a simple channel,

$$y = hx + n,$$

a basic equalizer may compute

$$\hat{x} = \frac{y}{\hat{h}}.$$

In practice, techniques such as MMSE equalization may provide improved robustness in noise.

1.14.4 Beamforming

With multiple antennas, the receiver can exploit spatial information.

The array output may be written as

$$z = \mathbf{w}^H \mathbf{y},$$

where $\mathbf{w}$ is a complex beamforming vector.

Properly selected weights can increase desired-signal gain while reducing interference.

Beamforming therefore adds a spatial dimension to the communication problem.

1.14.5 Interference Suppression

A receiver may estimate the structure of interfering signals and choose spatial, temporal, or frequency-domain processing that suppresses them.

In a multi-antenna system, this may involve identifying directions from which interference arrives and placing spatial nulls in those directions.

1.14.6 Adaptive Modulation and Coding

Wireless channel quality changes over time.

A robust communication system therefore adapts its transmission strategy.

When channel quality is poor, the transmitter may use

$$\text{lower-order modulation} + \text{stronger error correction.}$$

When channel quality improves, it may use

$$\text{higher-order modulation} + \text{higher coding rate.}$$

The objective is not necessarily to maximize instantaneous physical-layer rate.

Instead, the objective is to maximize useful delivered throughput:

$$\boxed{\text{goodput} = \text{successfully delivered useful data per unit time.}}$$

1.15 The Central Engineering Problem

Fixed wireless access can therefore be viewed as an optimization problem over a changing physical channel.

The system must continuously answer several questions:

- Which spatial beam should be used?
- What receive gain avoids clipping while preserving ADC resolution?
- What is the current channel response?
- How strong is the desired signal relative to interference and noise?
- Which modulation and coding scheme can the link currently support?
- When should a failed block be retransmitted?

These problems are tightly coupled.
For example,

beam selection

changes received signal power and interference.
That affects

SINR.

SINR affects

MCS selection.

The selected MCS affects

block error rate.

Block error rate affects

retransmission rate,

which finally affects

goodput.

Thus the complete link can be viewed as a closed-loop system:

$$\boxed{\text{channel} \rightarrow \text{measurement} \rightarrow \text{adaptation} \rightarrow \text{transmission} \rightarrow \text{feedback}}$$

This perspective is useful because modern wireless communication is not simply the process of transmitting a waveform from one antenna to another. It is an adaptive estimation and control problem in which the radio continually measures the environment and changes its behavior to maintain reliable and efficient communication.

The chapters that follow develop these concepts in greater detail, beginning with received-signal models, beamforming and acquisition, receiver dynamic range, SINR estimation, OFDM, link adaptation, and finally error correction and retransmission.

Chapter 2

The Wireless Signal Model

A wireless communication system ultimately attempts to recover information from an electromagnetic waveform that has been altered by propagation, interference, noise, and imperfections in the transmitter and receiver.

Before studying beamforming, OFDM, channel estimation, or link adaptation, it is useful to establish a mathematical model for this process.

The simplest useful model is

$$y = hx + i + n,$$

but each term in this equation represents a substantial physical process. This chapter develops the model gradually, beginning with a single transmitter and receiver and extending it to frequency-selective and multi-antenna channels.

2.1 From Transmitted Signal to Received Signal

Suppose a transmitter generates a signal $x(t)$. After the signal propagates through the environment, the receiver observes a modified waveform $y(t)$.

In the simplest idealized channel,

$$y(t) = h\,x(t),$$

where h represents the effect of propagation.

Real systems also contain noise and interference, giving

$$\boxed{y(t) = h\,x(t) + i(t) + n(t)}$$

where

- $x(t)$ is the desired transmitted signal,
- h is the propagation channel,
- $i(t)$ represents interfering signals,
- $n(t)$ represents thermal and receiver noise,
- $y(t)$ is the observed received signal.

This equation is one of the most important abstractions in communication theory.

Much of wireless receiver design can be understood as the problem of estimating $x(t)$ from $y(t)$ despite uncertainty in

$$h, \qquad i(t), \qquad n(t).$$

In other words,

Receiver design = recover x given an imperfect observation y.

2.2 Complex Baseband Representation

Radio signals are physically transmitted at a carrier frequency that may be hundreds of megahertz or several gigahertz.

A real RF waveform can be represented as

$$s_{\mathrm{RF}}(t) = A(t)\cos\left(2\pi f_c t + \phi(t)\right),$$

where

- $A(t)$ is the time-varying amplitude,
- f_c is the carrier frequency,
- $\phi(t)$ is the time-varying phase.

Working directly with this rapidly oscillating waveform is inconvenient.

Communication systems therefore commonly represent the signal using a complex baseband equivalent.

Define

$$x(t) = I(t) + jQ(t),$$

where

- $I(t)$ is the in-phase component,
- $Q(t)$ is the quadrature component,
- $j = \sqrt{-1}$.

The corresponding physical RF waveform can be written as

$$s_{\mathrm{RF}}(t) = \Re\left\{x(t)e^{j2\pi f_c t}\right\}.$$

Expanding,

$$s_{\mathrm{RF}}(t) = I(t)\cos(2\pi f_c t) - Q(t)\sin(2\pi f_c t).$$

This representation separates the rapidly oscillating carrier from the information-bearing baseband signal.

Consequently, most DSP algorithms operate on complex samples of the form

$$x[n] = I[n] + jQ[n].$$

This is why quantities throughout communication theory are frequently complex-valued.

2.3 The Complex Wireless Channel

The channel coefficient h is generally complex:

$$\boxed{h = |h|e^{j\theta}}$$

where

- $|h|$ represents amplitude scaling,
- θ represents phase rotation.

If

$$x = |x|e^{j\phi},$$

then

$$hx = |h||x|e^{j(\phi+\theta)}.$$

Thus the channel simultaneously changes the amplitude and phase of the transmitted signal. The received signal

$$y = hx$$

therefore becomes

$$|y| = |h||x|$$

and

$$\angle y = \angle x + \angle h.$$

The receiver must estimate these effects if it wishes to reconstruct the transmitted symbol.

2.4 Why Propagation Produces Phase Rotation

Consider a sinusoidal carrier with wavelength

$$\lambda = \frac{c}{f_c},$$

where c is the speed of light.
A propagation distance d creates a delay

$$\tau = \frac{d}{c}.$$

The corresponding carrier-phase shift is

$$\theta = -2\pi f_c \tau.$$

Substituting

$$\tau = \frac{d}{c}$$

gives

$$\theta = -\frac{2\pi d}{\lambda}.$$

Therefore a path-length change of one wavelength produces a full

$$2\pi$$

phase rotation.
A path-length change of

$$\frac{\lambda}{2}$$

produces approximately

$$\pi$$

radians, or 180°, of phase change.

At high carrier frequencies, wavelengths may be only a few centimeters. As a result, very small physical changes in the environment can produce substantial phase changes.

This observation becomes especially important in

- multipath propagation,
- beamforming,
- antenna arrays,
- coherent combining.

2.5 Path Loss

As an electromagnetic wave propagates away from a transmitter, its energy spreads over an increasingly large region.

In ideal free-space propagation, received power follows the Friis transmission equation:

$$P_r = P_t G_t G_r \left(\frac{\lambda}{4\pi d}\right)^2,$$

where

- P_t is transmitted power,
- P_r is received power,
- G_t is transmit antenna gain,
- G_r is receive antenna gain,
- d is propagation distance,

- λ is wavelength.

The free-space path loss is therefore

$$L_{\mathrm{FS}} = \left(\frac{4\pi d}{\lambda}\right)^2.$$

Expressed in decibels,

$$L_{\mathrm{FS,dB}} = 20\log_{10}\left(\frac{4\pi d}{\lambda}\right).$$

This immediately reveals two important effects.

Increasing distance increases path loss.

Increasing carrier frequency, and therefore decreasing wavelength, also increases free-space path loss for antennas of fixed gain.

Real propagation environments introduce additional effects beyond free-space spreading, including

- reflection,
- diffraction,
- scattering,
- absorption,
- shadowing.

These effects are incorporated into the channel coefficient h.

2.6 Noise

Even in the complete absence of external transmitters, a receiver observes random electrical fluctuations.

One important source is thermal noise.

The available thermal noise power over bandwidth B is approximately

$$P_N = kTB,$$

where

- k is Boltzmann's constant,
- T is absolute temperature,
- B is receiver bandwidth.

Thus,

larger bandwidth $\rightarrow$ more integrated noise power.

The receiver electronics themselves introduce additional noise.

Their degradation is often described using a noise figure.

In mathematical models, receiver noise is frequently approximated as complex additive white Gaussian noise:

$$n \sim \mathcal{CN}(0, \sigma_n^2).$$

This notation means that

$$\mathbb{E}[n] = 0$$

and

$$\mathbb{E}[|n|^2] = \sigma_n^2.$$

The Gaussian model is important because many independent random physical noise processes combine to produce approximately Gaussian behavior.

2.7 Signal-to-Noise Ratio

Suppose the received desired signal has power

$$P_S = \mathbb{E}\left[|hx|^2\right]$$

and the noise has power

$$P_N = \mathbb{E}\left[|n|^2\right].$$

The signal-to-noise ratio is

$$\boxed{\text{SNR} = \frac{P_S}{P_N}.}$$

In decibels,

$$\text{SNR}_{\text{dB}} = 10 \log_{10}\left(\frac{P_S}{P_N}\right).$$

Higher SNR generally allows the receiver to distinguish transmitted symbols more reliably.

However, practical wireless systems often operate in environments where interference is at least as important as thermal noise.

2.8 Interference

Suppose K other transmitters are active near the receiver.

A more complete model becomes

$$y = h_s x_s + \sum_{k=1}^{K} h_k x_k + n,$$

where

- x_s is the desired transmission,

- h_s is the desired channel,
- x_k is the signal from interferer k,
- h_k is the channel from interferer k.

The total interference is

$$i = \sum_{k=1}^{K} h_k x_k.$$

The model can therefore again be written compactly as

$$\boxed{y = h_s x_s + i + n.}$$

This decomposition is useful because the receiver's task is not simply to distinguish signal from random noise.

It must often distinguish one structured communication waveform from several other structured waveforms.

2.9 Signal-to-Interference-plus-Noise Ratio

If the desired signal power is P_S, total interference power is P_I, and noise power is P_N, then

$$\boxed{\text{SINR} = \frac{P_S}{P_I + P_N}.}$$

This quantity is usually more useful than SNR in interference-limited networks.

A receiver may observe a very strong total signal while still experiencing poor communication quality.

For example,

$$P_S = -50 \text{ dBm}$$

may appear strong.

However, if

$$P_I = -49 \text{ dBm},$$

the desired signal is competing with an interferer of comparable power.

The resulting SINR is poor even though the absolute received power is high.

This explains why received signal strength alone is generally insufficient for characterizing a communication link.

2.10 The Narrowband Channel Assumption

The simple model

$$y = hx + n$$

implicitly assumes that the entire transmitted signal experiences approximately the same channel coefficient h.

This is known as a *flat-fading* or narrowband approximation.

It is valid when the signal bandwidth is sufficiently small relative to the frequency scale over which the channel changes.

Under this assumption,

$$H(f) \approx h$$

throughout the signal bandwidth.

Every frequency component therefore experiences approximately the same amplitude scaling and phase rotation.

This assumption dramatically simplifies receiver design.

However, broadband communication systems often violate it.

2.11 Multipath Channel Model

In a realistic environment, the transmitted signal may reach the receiver through multiple paths.

Suppose there are L significant propagation paths.

The received waveform can be modeled as

$$y(t) = \sum_{\ell=0}^{L-1} h_\ell x(t - \tau_\ell) + n(t),$$

where

- h_ℓ is the complex gain of path ℓ,
- τ_ℓ is its propagation delay.

The corresponding channel impulse response is

$$\boxed{h(t) = \sum_{\ell=0}^{L-1} h_\ell \delta(t - \tau_\ell).}$$

The received waveform can then be expressed using convolution:

$$\boxed{y(t) = h(t) * x(t) + n(t).}$$

This is the natural generalization of the simple model

$$y = hx + n.$$

For a single propagation path,

$$h(t) = h_0 \delta(t - \tau_0),$$

and the multipath model reduces approximately to a scaled and delayed version of the original signal.

2.12 Frequency-Domain Channel Representation

Convolution in time corresponds to multiplication in frequency.

Taking the Fourier transform of

$$y(t) = h(t) * x(t) + n(t)$$

produces

$$\boxed{Y(f) = H(f)X(f) + N(f).}$$

For the multipath channel,

$$H(f) = \sum_{\ell=0}^{L-1} h_\ell e^{-j2\pi f \tau_\ell}.$$

This equation is extremely important.

Each propagation path contributes a complex rotating term

$$h_\ell e^{-j2\pi f \tau_\ell}.$$

Because the phase rotation depends on frequency, the paths may add constructively at some frequencies and destructively at others.

Thus,

$$|H(f)|$$

may vary significantly across the signal bandwidth.

This is known as *frequency-selective fading.*

2.13 An Example of Frequency-Selective Fading

Consider a simple two-path channel:

$$H(f) = h_0 + h_1 e^{-j2\pi f \tau}.$$

Suppose the paths have similar amplitudes:

$$h_0 \approx h_1.$$

At frequencies where

$$2\pi f \tau \approx 2\pi m,$$

for integer m, the two paths have similar phase and reinforce one another.

At frequencies where

$$2\pi f \tau \approx (2m+1)\pi,$$

the two paths are approximately opposite in phase and may largely cancel.

Thus the same physical channel can contain frequencies with strong gain and frequencies with deep attenuation.

These strongly attenuated frequencies are often called *deep fades*.

Conceptually,

multipath → frequency-dependent phase → constructive/destructive interference → frequency-selective fading.

2.14 Delay Spread

The difference between the earliest and latest significant multipath arrivals is related to the channel's delay spread.

For example, if the first significant path arrives at

$$\tau_{\min}$$

and the last at

$$\tau_{\max},$$

then a rough delay-spread measure is

$$T_{\text{delay}} \approx \tau_{\max} - \tau_{\min}.$$

Large delay spread implies that delayed copies of one transmitted symbol may overlap neighboring symbols.

This produces *intersymbol interference*, or ISI.

Thus,

$$\boxed{\text{multipath delay} \rightarrow \text{symbol overlap} \rightarrow \text{ISI.}}$$

One of the major motivations for OFDM is to make this wideband multipath problem substantially easier to handle.

2.15 Discrete-Time Channel Model

Digital receivers operate on sampled signals.

After sampling, the multipath channel may be represented as

$$y[n] = \sum_{\ell=0}^{L-1} h[\ell]x[n-\ell] + n[n].$$

Equivalently,

$$y[n] = h[n] * x[n] + n[n].$$

The sequence

$$h[0], h[1], \ldots, h[L-1]$$

contains the discrete-time channel taps.

Each tap describes the contribution of a delayed version of the transmitted signal.

This representation is common in practical DSP implementations because channel estimation often amounts to estimating these taps or their frequency-domain equivalents.

2.16 Time-Varying Channels

The channel may also change with time.

A more complete representation is

$$h(t, \tau),$$

where

- t describes when the channel is observed,
- τ represents propagation delay.

The received signal becomes

$$y(t) = \int h(t, \tau) x(t - \tau)\, d\tau + n(t).$$

Even in fixed wireless systems, $h(t, \tau)$ may change because of

- moving vehicles,
- foliage,
- people,
- changing reflectors,
- temperature-dependent hardware drift,
- oscillator drift.

Thus a receiver often estimates the channel repeatedly rather than assuming that one estimate remains valid indefinitely.

2.17 Channel Coherence

A useful practical concept is the *coherence time*.

The coherence time describes approximately how long the channel remains sufficiently similar that a previous channel estimate can still be considered valid.

Conceptually,

$$T_{\text{estimate}} \ll T_{\text{coherence}}$$

is desirable.

If channel estimation happens too slowly relative to channel variation, then

$$\hat{h}$$

may already be stale by the time it is used.

There is an analogous frequency-domain quantity called *coherence bandwidth*, which describes the frequency range over which the channel remains approximately similar.

A signal bandwidth much smaller than the coherence bandwidth experiences approximately flat fading.

A signal bandwidth much larger than the coherence bandwidth experiences frequency-selective fading.

2.18 Channel Estimation

The receiver usually does not know h in advance.

It must estimate it from the received signal.

Suppose the transmitter sends a known pilot symbol x_p.

The receiver observes

$$y_p = hx_p + n.$$

Ignoring noise temporarily,

$$h = \frac{y_p}{x_p}.$$

Thus a simple channel estimate is

$$\boxed{\hat{h} = \frac{y_p}{x_p}.}$$

With multiple pilot observations, the receiver can average information to reduce estimation noise.

For example,

$$\hat{h} = \frac{\sum_{k=1}^{N} x_k^* y_k}{\sum_{k=1}^{N} |x_k|^2}.$$

This is closely related to a least-squares estimate.

Channel estimation is fundamental because many subsequent receiver operations depend on

$$\hat{h}.$$

These include

- equalization,
- coherent demodulation,
- SINR estimation,
- beamforming,
- link adaptation.

2.19 Equalization

Suppose the channel is narrowband and

$$y = hx + n.$$

If the receiver knows h, a simple zero-forcing equalizer computes

$$\hat{x} = \frac{y}{h}.$$

Substituting the signal model,

$$\hat{x} = \frac{hx + n}{h},$$

giving

$$\hat{x} = x + \frac{n}{h}.$$

The channel distortion has been removed.

However, there is an important problem.

If

$$|h| \approx 0,$$

then

$$\frac{n}{h}$$

can become extremely large.

Thus blindly inverting the channel can significantly amplify noise in deep fades.

More sophisticated equalizers, such as the Minimum Mean-Squared Error (MMSE) equalizer, explicitly balance channel inversion against noise amplification.

This illustrates a recurring theme in communication systems:

perfectly undoing one impairment may amplify another.

2.20 Extension to Multiple Receive Antennas

Suppose the receiver contains N antennas.

Instead of observing one scalar sample y, the receiver observes a vector:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}.$$

For one desired transmitter,

$$\boxed{\mathbf{y} = \mathbf{h}x + \mathbf{n}}$$

where

$$\mathbf{h} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_N \end{bmatrix}.$$

Each coefficient h_k represents the channel from the transmitter to receive antenna k.

Because each antenna occupies a different physical position,

$$h_1, h_2, \ldots, h_N$$

generally have different phases and possibly different amplitudes.

The vector

$$\mathbf{h}$$

therefore contains spatial information about the incoming wave.

This is the foundation of beamforming.

2.21 Phase Across an Antenna Array

Consider two antennas separated by distance d.

Suppose a plane wave arrives at angle θ.

The additional path traveled before reaching one antenna relative to the other is approximately

$$\Delta r = d \sin \theta.$$

The corresponding phase difference is

$$\Delta \phi = \frac{2\pi}{\lambda} d \sin \theta.$$

For a uniform linear array with N elements, the spatial phase progression can therefore be represented by a steering vector such as

$$\mathbf{a}(\theta) = \begin{bmatrix} 1 \\ e^{-j\psi} \\ e^{-j2\psi} \\ \vdots \\ e^{-j(N-1)\psi} \end{bmatrix},$$

where

$$\psi = \frac{2\pi d}{\lambda} \sin \theta.$$

This phase pattern contains information about the direction from which the signal arrived.

Beamforming exploits precisely this relationship.

2.22 Multi-Antenna Signal Model with Interference

Suppose there is one desired transmitter and K interferers.

The received antenna vector becomes

$$\boxed{\mathbf{y} = \mathbf{h}_s x_s + \sum_{j=1}^{K} \mathbf{h}_j x_j + \mathbf{n}}$$

where

- $\mathbf{h}_s$ is the spatial channel of the desired transmitter,
- $\mathbf{h}_j$ is the spatial channel of interferer j,

- x_s is the desired transmitted symbol,
- x_j is the signal transmitted by interferer j,
- $\mathbf{n}$ is the receiver-noise vector.

The desired channel is

$$\mathbf{h}_s = \begin{bmatrix} h_{s,1} \\ h_{s,2} \\ \vdots \\ h_{s,N} \end{bmatrix}.$$

An interferer arriving from a different direction generally produces a different channel vector:

$$\mathbf{h}_j \neq \mathbf{h}_s.$$

This difference creates the possibility of spatial separation.

2.23 The Beamforming Vector

A beamformer combines the antenna signals using complex weights.

Let

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix}$$

with

$$w_k = a_k e^{j\phi_k}.$$

Thus,

$$\boxed{\mathbf{w} = \begin{bmatrix} a_1 e^{j\phi_1} \\ a_2 e^{j\phi_2} \\ \vdots \\ a_N e^{j\phi_N} \end{bmatrix}.}$$

Each weight contains both

- an amplitude a_k,
- a phase ϕ_k.

The beamformer output is

$$\boxed{z = \mathbf{w}^H \mathbf{y},}$$

where

$$\mathbf{w}^H$$

is the Hermitian transpose, or conjugate transpose, of $\mathbf{w}$.

Explicitly,

$$z = w_1^* y_1 + w_2^* y_2 + \cdots + w_N^* y_N.$$

Thus the beamformer phase-shifts and scales each antenna signal before summing them.

2.24 Coherent Combining

Suppose the desired channel is

$$\mathbf{h} = \begin{bmatrix} |h_1| e^{j\theta_1} \\ |h_2| e^{j\theta_2} \\ \vdots \\ |h_N| e^{j\theta_N} \end{bmatrix}.$$

Without phase correction, the antenna signals may partially cancel.

The beamformer can compensate for these phases by choosing weights that approximately undo them.

For example,

$$w_k \propto h_k$$

gives

$$w_k^* h_k \propto |h_k|^2.$$

The individual antenna contributions then add constructively.

This is called *coherent combining*.

The central idea is

$$\boxed{\text{align the desired phases before summing.}}$$

2.25 Maximum-Ratio Combining

For independent equal-power receiver noise and no significant interference, a particularly important beamforming solution is Maximum-Ratio Combining, or MRC.

The beamformer is chosen proportional to the desired channel:

$$\boxed{\mathbf{w} \propto \mathbf{h}_s.}$$

A normalized form is

$$\mathbf{w} = \frac{\mathbf{h}_s}{\|\mathbf{h}_s\|}.$$

The output desired signal becomes

$$z_s = \mathbf{w}^H \mathbf{h}_s x_s.$$

Substituting the normalized MRC weights,

$$z_s = \frac{\mathbf{h}_s^H \mathbf{h}_s}{\|\mathbf{h}_s\|} x_s.$$

Since

$$\mathbf{h}_s^H \mathbf{h}_s = \|\mathbf{h}_s\|^2,$$

we obtain

$$z_s = \|\mathbf{h}_s\| x_s.$$

The array has coherently combined the desired signal from all antennas.

2.26 Desired Signal Power After Beamforming

The desired component of the beamformer output is

$$z_s = \mathbf{w}^H \mathbf{h}_s x_s.$$

If the transmitted symbol has power

$$P_x = \mathbb{E}[|x_s|^2],$$

then the desired output power is

$$\boxed{P_s = P_x \left| \mathbf{w}^H \mathbf{h}_s \right|^2 .}$$

This equation is central to beamforming.
The quantity

$$\left| \mathbf{w}^H \mathbf{h}_s \right|^2$$

is the effective spatial gain applied to the desired channel.

Beamforming therefore selects $\mathbf{w}$ to control how strongly signals from different spatial channels appear at the output.

2.27 Interference After Beamforming

For interferer j, the output contribution is

$$z_j = \mathbf{w}^H \mathbf{h}_j x_j.$$

Its output power is

$$P_{I,j} = P_j \left| \mathbf{w}^H \mathbf{h}_j \right|^2 ,$$

where

$$P_j = \mathbb{E}[|x_j|^2].$$

Therefore changing $\mathbf{w}$ changes both the desired signal and the interference.

This is crucial.

A beamformer should not necessarily maximize

$$\left|\mathbf{w}^H \mathbf{h}_s\right|^2$$

alone.

It may be more useful to sacrifice a small amount of desired-signal gain if doing so substantially suppresses interference.

2.28 Spatial Nulls

Suppose the beamformer can choose weights satisfying

$$\mathbf{w}^H \mathbf{h}_j \approx 0$$

for a particular interferer.

Then

$$z_j \approx 0.$$

The beamformer has created a *spatial null* in the direction or spatial channel associated with that interferer.

At the same time, we would like

$$\left|\mathbf{w}^H \mathbf{h}_s\right|$$

to remain large.

Thus beamforming can be viewed as the simultaneous construction of

gain toward desired signals

and

nulls toward unwanted signals.

This is one of the principal advantages of multi-antenna systems.

2.29 Noise After Beamforming

Suppose receiver noise has covariance matrix

$$\mathbf{R}_n = \mathbb{E}\left[\mathbf{n}\mathbf{n}^H\right].$$

The beamformer output noise is

$$n_{\text{out}} = \mathbf{w}^H \mathbf{n}.$$

Its power is

$$\boxed{P_{N,\text{out}} = \mathbf{w}^H \mathbf{R}_n \mathbf{w}.}$$

If the antenna noises are independent and have equal variance,

$$\mathbf{R}_n = \sigma_n^2 \mathbf{I}.$$

Then

$$P_{N,\text{out}} = \sigma_n^2 \mathbf{w}^H \mathbf{w}.$$

If the beamformer is normalized such that

$$\|\mathbf{w}\|^2 = 1,$$

then

$$P_{N,\text{out}} = \sigma_n^2.$$

2.30 Interference Covariance

Rather than representing each interferer separately, it is often useful to describe their combined spatial statistics using a covariance matrix.

Define

$$\mathbf{R}_{I+N} = \mathbb{E}\left[(\mathbf{i}+\mathbf{n})(\mathbf{i}+\mathbf{n})^H\right].$$

This matrix contains both

- the total interference power,
- information about how the interference is distributed spatially across the antenna array.

The interference-plus-noise power after beamforming is

$$\boxed{P_{I+N} = \mathbf{w}^H \mathbf{R}_{I+N} \mathbf{w}.}$$

This quantity becomes central when designing interference-aware beamformers.

2.31 Post-Beamforming SINR

The desired output power is

$$P_S = P_x \left|\mathbf{w}^H \mathbf{h}_s\right|^2.$$

The interference-plus-noise output power is

$$P_{I+N} = \mathbf{w}^H \mathbf{R}_{I+N} \mathbf{w}.$$

Therefore,

$$\boxed{\mathrm{SINR}(\mathbf{w}) = \frac{P_x \left|\mathbf{w}^H \mathbf{h}_s\right|^2}{\mathbf{w}^H \mathbf{R}_{I+N} \mathbf{w}}.}$$

This equation provides a much deeper interpretation of beamforming than simply "pointing an antenna."

The weights influence both numerator and denominator.

A beam may slightly reduce desired-signal gain while dramatically reducing interference.

In such a case,

$$P_S \downarrow$$

slightly,

while

$$P_I \downarrow$$

substantially.

The resulting SINR may increase.

Thus,

$$\boxed{\text{maximum received power} \neq \text{maximum communication quality.}}$$

2.32 Beamforming as an Optimization Problem

The previous equation suggests a natural mathematical objective:

$$\boxed{\mathbf{w}^* = \arg\max_{\mathbf{w}} \frac{P_x |\mathbf{w}^H \mathbf{h}_s|^2}{\mathbf{w}^H \mathbf{R}_{I+N} \mathbf{w}}.}$$

The exact solution depends on what information is available and what constraints are imposed on the weights.

For example, a radio may impose

- constant-amplitude weights,
- discrete phase settings,
- maximum per-antenna power,
- finite-resolution phase control,
- predefined codebook beams.

Thus the mathematically optimal solution may differ from the solution that is practical in real hardware.

This distinction between theoretical optimum and hardware-constrained implementation appears repeatedly in real communication systems.

2.33 Transmit Beamforming

The same basic concept can be applied at the transmitter.

Suppose one symbol x is transmitted through N antennas with weight vector

$$\mathbf{w}_t.$$

The transmitted antenna signals are

$$\mathbf{x} = \mathbf{w}_t x.$$

If the channel from the transmit array to a single receiver is $\mathbf{h}$, the received signal becomes

$$y = \mathbf{h}^H \mathbf{w}_t x + n.$$

The effective channel is therefore

$$h_{\text{eff}} = \mathbf{h}^H \mathbf{w}_t.$$

By choosing the phases of $\mathbf{w}_t$, the transmitter can cause signals emitted by different antennas to combine constructively at the receiver.

Thus receive beamforming performs

$$\text{weighted combining after propagation,}$$

whereas transmit beamforming performs

$$\text{weighted transmission before propagation.}$$

Both exploit the same underlying spatial phase relationships.

2.34 Multiple Transmit and Receive Antennas

If there are N_t transmit antennas and N_r receive antennas, the channel becomes a matrix:

$$\boxed{\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}.}$$

The matrix

$$\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$$

contains the channel between every transmit and receive antenna pair.

Explicitly,

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1N_t} \\ h_{21} & h_{22} & \cdots & h_{2N_t} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_r1} & h_{N_r2} & \cdots & h_{N_rN_t} \end{bmatrix}.$$

Here,

$$h_{ij}$$

represents the channel from transmit antenna j to receive antenna i.

This matrix formulation forms the basis of Multiple-Input Multiple-Output, or MIMO, communication.

MIMO systems can use their spatial degrees of freedom for different purposes, including

- beamforming,
- diversity,
- interference suppression,
- spatial multiplexing.

2.35 Beamforming Versus Spatial Multiplexing

These two concepts should be distinguished.

In beamforming, multiple antennas cooperate to improve transmission of essentially the same data stream.

Conceptually,

$$\text{many antennas} \rightarrow \text{one strengthened spatial stream.}$$

In spatial multiplexing, the transmitter attempts to send multiple independent data streams simultaneously:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix}.$$

The receiver uses the channel matrix

$$\mathbf{H}$$

to separate them.

Conceptually,

$$\text{many antennas} \rightarrow \text{multiple simultaneous information streams.}$$

Beamforming primarily seeks improved link quality or spatial selectivity.

Spatial multiplexing primarily seeks increased throughput.

A practical MIMO system may use combinations of both.

2.36 The Wireless Channel as an Estimation Problem

The mathematical models developed in this chapter reveal an important property of practical wireless communication.

The receiver rarely knows exactly

$$h, \qquad H(f), \qquad \mathbf{h}, \qquad \mathbf{H},$$

or

$$\mathbf{R}_{I+N}.$$

These quantities must be estimated from noisy observations.

Thus many receiver algorithms follow the structure

$$\boxed{\text{observe} \rightarrow \text{estimate} \rightarrow \text{compensate} \rightarrow \text{decode.}}$$

For example,

$$\text{pilot observations} \rightarrow \hat{H}[k] \rightarrow \text{equalization},$$

or

$$\text{antenna observations} \rightarrow \hat{\mathbf{h}} \rightarrow \text{beamforming}.$$

The quality of these estimates directly affects communication performance.

2.37 The Wireless Channel as a Control Problem

Modern radio systems go one step further.

They do not merely estimate the channel; they change their behavior in response to it.

Measurements may influence

$$\text{beam direction},$$

$$\text{receiver gain},$$

$$\text{transmit power},$$

$$\text{modulation},$$

$$\text{coding rate},$$

and

$$\text{retransmission behavior}.$$

The communication system therefore forms a feedback loop:

$$\boxed{\text{channel} \rightarrow \text{measurement} \rightarrow \text{estimate} \rightarrow \text{decision} \rightarrow \text{transmission} \rightarrow \text{channel.}}$$

This perspective is particularly useful for understanding high-performance fixed wireless systems.

The physical channel is not merely something the receiver passively tolerates.

It is continuously measured, estimated, and exploited.

2.38 From the Signal Model to Beam Acquisition

The multi-antenna model

$$\mathbf{y} = \mathbf{h}_s x_s + \sum_j \mathbf{h}_j x_j + \mathbf{n}$$

assumes that the receiver already has some knowledge of

$$\mathbf{h}_s.$$

During initial link acquisition, however, this information may not yet be available. The system may not know

- which spatial direction contains the desired transmitter,
- which beam maximizes useful signal quality,
- what receive gain should be used,
- what timing offset exists,
- what frequency offset exists,
- how much interference is present.

The receiver must therefore search the spatial domain and identify a usable link. This leads naturally to the next problem:

How do two radios discover and acquire the correct beam?

The next chapter develops beam acquisition and beam search, including the practical interaction between beam selection, signal strength, interference, gain control, and measurement reliability.

Chapter 3

Beam Acquisition and RN–BN Lookup

Beam acquisition is the process by which two radios discover a usable spatial communication path and determine how their antenna arrays should be configured to communicate reliably.

In a fixed wireless system, this often occurs when a Remote Node (RN) first attempts to establish communication with a Base Node (BN), or when an existing link must be reacquired after interruption.

At first glance, this might appear to be a simple directional-search problem: point the antennas toward one another and communicate. In practice, however, the problem is substantially more complicated.

The radio may not initially know

- the direction from which the desired signal will arrive,
- the best transmit beam,
- the best receive beam,
- the propagation channel,
- the interference environment,
- the appropriate receive gain,
- the timing offset,
- the carrier-frequency offset.

Furthermore, the best communication path may not correspond to the geometric line between the two radios.

Reflections from buildings, terrain, and other structures can produce strong alternative propagation paths.

Beam acquisition is therefore fundamentally a problem of

$$\boxed{\text{search} + \text{measurement} + \text{estimation} + \text{decision}.}$$

This chapter develops the fundamental ideas behind that process.

3.1 Why Beam Acquisition Is Necessary

An omnidirectional antenna radiates or receives energy over a broad angular region.

This makes discovery relatively easy because the radio does not need to know where the other device is located.

However, spreading electromagnetic energy over many directions reduces the amount of power delivered in any one direction.

A directional antenna concentrates energy spatially.

Conceptually,

$$\text{omnidirectional transmission} \rightarrow \text{energy spread over many directions},$$

whereas

$$\text{directional transmission} \rightarrow \text{energy concentrated into a smaller angular region}.$$

This concentration produces antenna gain.

Directional communication can therefore improve

- communication range,
- received signal strength,
- SINR,
- interference rejection,
- achievable modulation order,
- throughput.

The price of this gain is that the radio must know approximately where to direct the beam.

Thus,

$$\boxed{\text{greater spatial gain} \rightarrow \text{greater need for spatial knowledge.}}$$

Beam acquisition exists to obtain that knowledge.

3.2 From Antennas to Beams

A beam is created when signals from multiple antenna elements are combined with carefully chosen phase and amplitude relationships.

Consider a uniform linear array containing N antennas separated by distance d.

Suppose a plane wave arrives from angle θ.

The additional path length between neighboring antennas is approximately

$$\Delta r = d \sin\theta.$$

The corresponding phase difference is

$$\Delta\phi = \frac{2\pi d}{\lambda} \sin\theta,$$

where

$$\lambda = \frac{c}{f_c}$$

is the carrier wavelength.

The received signal therefore has a predictable phase progression across the array.

A simplified steering vector is

$$\mathbf{a}(\theta) = \begin{bmatrix} 1 \\ e^{-j\psi} \\ e^{-j2\psi} \\ \vdots \\ e^{-j(N-1)\psi} \end{bmatrix},$$

where

$$\psi = \frac{2\pi d}{\lambda} \sin\theta.$$

The array therefore contains information about the direction of arrival through relative phase.

3.3 Beamforming as Spatial Filtering

Suppose the antenna array receives the vector

$$\mathbf{y}.$$

A beamformer applies a complex weight vector

$$\mathbf{w}$$

and forms

$$z = \mathbf{w}^H \mathbf{y}.$$

If the beamforming vector is chosen to align with signals arriving from a particular direction, those signals add coherently.

Signals arriving from different directions generally do not align as well and therefore contribute less strongly.

The beamformer can therefore be interpreted as a spatial filter.

Ordinary filters separate signals according to frequency.

Beamformers separate signals according to spatial structure.

Conceptually,

$$\boxed{\text{frequency filter} \rightarrow \text{selects frequency}}$$

while

$$\boxed{\text{beamformer} \rightarrow \text{selects spatial direction or channel.}}$$

This is an important interpretation.

A beam is not merely a geometric object pointing in one direction. It is the spatial response produced by a particular set of antenna weights.

3.4 The Array Response

Suppose a signal arrives with steering vector

$$\mathbf{a}(\theta).$$

The response of beamformer $\mathbf{w}$ to that signal is

$$B(\theta) = \mathbf{w}^H \mathbf{a}(\theta).$$

The corresponding power response is

$$\boxed{G(\theta) = \left|\mathbf{w}^H \mathbf{a}(\theta)\right|^2.}$$

The function

$$G(\theta)$$

describes the beam pattern.
At angles where

$$G(\theta)$$

is large, the array has high sensitivity.
At angles where

$$G(\theta)$$

is small, the array suppresses incoming signals.
A beam pattern generally contains

- a main lobe,
- side lobes,
- possibly spatial nulls.

The main lobe describes the principal direction of sensitivity.
Side lobes represent smaller regions of unintended gain.
Nulls are directions where the array response becomes very small.

3.5 Beam Width and Array Gain

One of the most fundamental beamforming tradeoffs is between beam width and gain.

A wide beam covers a large angular region.

This makes acquisition easier because the remote radio can be detected over a broad range of directions.

However, energy is distributed across that region, so the peak spatial gain is relatively low.

A narrow beam concentrates energy more strongly.

This increases gain but covers a smaller angular region.

Conceptually,

$$\boxed{\text{wide beam} \rightarrow \text{large angular coverage} \rightarrow \text{lower gain}}$$

while

$$\boxed{\text{narrow beam} \rightarrow \text{small angular coverage} \rightarrow \text{higher gain.}}$$

Larger antenna arrays can generally produce narrower beams.

For a simple array, beam width scales approximately inversely with aperture size:

$$\theta_{\text{BW}} \propto \frac{\lambda}{D},$$

where

$$D$$

is the physical aperture of the array.

Thus,

$$D \uparrow \quad \Rightarrow \quad \theta_{\text{BW}} \downarrow .$$

This provides greater directional resolution but also increases acquisition complexity.

3.6 Beam Codebooks

Although a beamformer can theoretically use an enormous number of possible weight combinations, practical radios often restrict operation to a predefined set of beamforming vectors.

This set is called a *beam codebook*.

Suppose the codebook contains

$$\mathcal{W} = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{N_b}\}.$$

Each vector corresponds to a different spatial beam.

Instead of solving continuously for arbitrary weights during acquisition, the radio can test

$$\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{N_b}$$

and determine which provides the best link.

This converts beam acquisition into a discrete search problem.

For example,

$$\boxed{\text{Which } \mathbf{w}_k \text{ produces the best communication quality?}}$$

The answer may depend on the propagation environment as well as the physical direction of the remote radio.

3.7 The RN–BN Acquisition Problem

Consider a Base Node with a transmit-beam codebook containing

$$N_t$$

candidate beams.
Suppose the Remote Node has

$$N_r$$

candidate receive beams.
A particular link configuration is described by the pair

$$(\mathbf{w}_t, \mathbf{w}_r).$$

If every transmit beam is tested against every receive beam, the total number of combinations is

$$\boxed{N_{\text{pairs}} = N_t N_r.}$$

For example, if

$$N_t = 32$$

and

$$N_r = 16,$$

then

$$N_{\text{pairs}} = 512.$$

The system must somehow identify a useful configuration among these possibilities.

3.8 Exhaustive Beam Search

The most direct strategy is exhaustive search.
For each transmit beam

$$\mathbf{w}_{t,i},$$

the receiver tests every receive beam

$$\mathbf{w}_{r,j}.$$

The system measures some quality metric

$$M_{ij}.$$

It then selects

$$(i^*, j^*) = \arg\max_{i,j} M_{ij}.$$

If each beam pair requires measurement time

$$T_m,$$

then total search time is approximately

$$\boxed{T_{\text{search}} = N_t N_r T_m.}$$

Exhaustive search has an important advantage:

it is difficult to miss the best codebook pair.

However, it can be slow when the codebooks are large.

For example, increasing the number of beams improves angular resolution but also increases the number of measurements required.

Thus,

$$\boxed{\text{angular resolution} \leftrightarrow \text{search complexity.}}$$

3.9 Why Search Time Matters

Acquisition time affects both user experience and network efficiency.

During acquisition, radio resources may be consumed without delivering substantial user data.

Long searches can increase

- initial connection latency,
- recovery time after link loss,
- signaling overhead,
- energy consumption.

If the channel itself changes during the search, another problem appears.

Suppose

$$T_{\text{search}}$$

becomes comparable with the channel coherence time

$$T_c.$$

Then the conditions measured at the beginning of the search may no longer match those at the end.

Ideally,

$$\boxed{T_{\text{search}} \ll T_c.}$$

This motivates faster search methods.

3.10 Hierarchical Beam Search

A common alternative to exhaustive search is hierarchical or coarse-to-fine acquisition.

The basic idea is to begin with a small number of wide beams.

Suppose the entire angular region is first divided into four broad sectors.

The system tests those beams and identifies the strongest candidate region.

It then subdivides that region into narrower beams.

The process may continue through several levels:

$$\boxed{\text{wide beam} \rightarrow \text{medium beam} \rightarrow \text{narrow beam.}}$$

This resembles a tree search.

For example,

$$\text{entire field of view}$$

may first be divided into

$$4$$

regions.

The best region may then be divided into another

$$4$$

subregions.

Instead of measuring every possible narrow beam, the system concentrates measurements where the signal appears most promising.

3.11 The Fundamental Hierarchical-Search Tradeoff

Hierarchical acquisition reduces search time but introduces risk.

Wide beams have lower gain.

A weak signal that would be detectable using a narrow high-gain beam may be difficult to detect using the initial wide beam.

Thus,

$$\boxed{\text{faster search} \leftrightarrow \text{reduced initial beam gain.}}$$

There is also the possibility that noise or interference causes the coarse stage to choose the wrong region.

Once the algorithm enters the wrong branch of the search tree, it may never test the true optimal narrow beam.

Practical hierarchical acquisition therefore often retains multiple candidates rather than only the single strongest beam.

3.12 Beam Sweeping

A common acquisition mechanism is *beam sweeping*.

One radio systematically transmits known signals through a sequence of beams:

$$\mathbf{w}_{t,1}, \mathbf{w}_{t,2}, \ldots, \mathbf{w}_{t,N_t}.$$

The receiver observes each transmission and records a quality metric.

For example,

$$M_1, M_2, \ldots, M_{N_t}.$$

It can then identify promising transmit beams.

The roles may then be reversed or the receive beam may be swept.

Conceptually,

$$\boxed{\text{transmit known signal} \rightarrow \text{sweep beams} \rightarrow \text{measure response} \rightarrow \text{choose candidate.}}$$

Known signals are important because the receiver must distinguish the desired transmitter from arbitrary RF energy.

3.13 Pilot and Preamble Detection

During initial acquisition, the receiver does not yet have ordinary user data available for channel estimation.

The transmitter therefore sends a known sequence, often called a

- pilot,
- preamble,
- synchronization sequence,
- reference signal.

Suppose the known sequence is

$$x[0], x[1], \ldots, x[N-1].$$

The receiver observes

$$y[n].$$

One simple detection mechanism is correlation:

$$C(\tau) = \sum_{n=0}^{N-1} y[n+\tau] x^*[n].$$

If the known sequence begins near delay

$$\tau_0,$$

then

$$|C(\tau)|$$

should have a peak near

$$\tau = \tau_0.$$

Thus correlation helps answer two questions:

1. Is the desired transmitter present?
2. Where in time does its known sequence begin?

This means that acquisition often combines spatial and temporal search.

3.14 Spatial and Temporal Acquisition

Beam acquisition cannot always be separated cleanly from synchronization.
The receiver may simultaneously be uncertain about

$$\text{beam}$$

and

$$\text{timing}.$$

A conceptual search may therefore involve

$$M(i, \tau),$$

where

$$i$$

indexes beam configuration and

$$\tau$$

indexes timing hypothesis.
The receiver may look for

$$(i^*, \tau^*) = \arg\max_{i,\tau} M(i, \tau).$$

Carrier-frequency offset may add another dimension:

$$M(i, \tau, \Delta f).$$

This demonstrates why initial acquisition can be computationally expensive.
The radio may effectively be searching through

$$\boxed{\text{space} \times \text{time} \times \text{frequency}.}$$

3.15 Beam-Scoring Metrics

Once a candidate beam has been tested, the receiver must assign it some measure of quality.

Possible metrics include

- received power,
- RSSI,
- pilot correlation magnitude,
- SNR,
- SINR,
- EVM,
- packet-detection probability,
- decoder success rate.

Different metrics provide different information.

No single metric is universally optimal for every phase of acquisition.

3.16 Received Power and RSSI

The simplest measurement is total received power.

If a beam produces

$$P_{\mathrm{RX}},$$

then the system may initially prefer beams with larger values.

This is inexpensive and fast to compute.

However, received power includes more than the desired transmitter.

Approximately,

$$P_{\mathrm{RX}} = P_S + P_I + P_N.$$

Therefore a strong interfering transmitter may cause a beam to appear attractive even when communication quality is poor.

Thus,

$$\boxed{\text{high RSSI} \not\Rightarrow \text{high-quality link.}}$$

3.17 Pilot Correlation

Correlation with a known sequence provides stronger evidence that received energy belongs to the desired transmitter.

A normalized metric might take the form

$$M_{\mathrm{corr}} = \frac{\left|\sum_n y[n]x^*[n]\right|^2}{\sum_n |y[n]|^2}.$$

A high value indicates that the received waveform resembles the expected pilot.
This is particularly useful during initial discovery because it distinguishes

$$\text{desired structured signal}$$

from

$$\text{unrelated RF energy.}$$

3.18 SNR and SINR

Once the desired signal can be identified, a more useful metric is often signal quality.
The SNR is

$$\text{SNR} = \frac{P_S}{P_N},$$

while SINR is

$$\boxed{\text{SINR} = \frac{P_S}{P_I + P_N}.}$$

SINR is particularly useful in environments with strong interference.
Consider two beams.
Beam A produces

$$P_S = -55 \text{ dBm}$$

and

$$P_I + P_N = -85 \text{ dBm}.$$

Its SINR is approximately

$$30 \text{ dB}.$$

Beam B produces a stronger desired signal:

$$P_S = -50 \text{ dBm},$$

but also strong interference:

$$P_I + P_N = -52 \text{ dBm}.$$

Its SINR is approximately

$$2 \text{ dB}.$$

Beam B has greater desired-signal power, but Beam A provides the much better communication link.
This illustrates a fundamental principle:

$$\boxed{\text{the strongest beam is not always the best beam.}}$$

3.19 EVM and Decoder-Based Metrics

Once synchronization and channel estimation are sufficiently reliable, the system may evaluate beam quality using demodulation performance.

For known symbols,

$$\mathrm{EVM}^2 = \frac{\mathbb{E}\left[|\hat{X} - X|^2\right]}{\mathbb{E}\left[|X|^2\right]}.$$

Smaller EVM generally indicates a cleaner received signal.

At an even higher layer, the system can observe whether transmitted blocks actually decode successfully.

Possible metrics include

$$\text{BLER},$$

ACK/NACK statistics, or decoder confidence.

These metrics are very meaningful but generally require more observation time than simple power measurements.

Thus there is often a hierarchy:

$$\boxed{\text{power} \rightarrow \text{correlation} \rightarrow \text{SINR/EVM} \rightarrow \text{decoder performance.}}$$

Each stage becomes more informative but usually more expensive.

3.20 Measurement Dwell Time

A beam cannot usually be evaluated from a single noisy sample.

Suppose the radio measures power samples

$$P_1, P_2, \ldots, P_N.$$

An estimate of average power is

$$\hat{P} = \frac{1}{N} \sum_{k=1}^{N} P_k.$$

As the number of samples increases, the estimate generally becomes more stable.

However, the measurement requires more time.

If the sampling interval is

$$T_s,$$

then the dwell time is approximately

$$T_{\text{dwell}} = NT_s.$$

Thus,

$$N \uparrow \quad \Rightarrow \quad \text{measurement variance} \downarrow$$

but also

$$N \uparrow \quad \Rightarrow \quad T_{\text{dwell}} \uparrow .$$

This gives another fundamental acquisition tradeoff:

$$\boxed{\text{measurement reliability} \leftrightarrow \text{measurement speed.}}$$

3.21 Averaging

One way to improve measurement reliability is temporal averaging.

Suppose the instantaneous metric is

$$M_k.$$

A simple moving average is

$$\bar{M}_k = \frac{1}{L} \sum_{\ell=0}^{L-1} M_{k-\ell}.$$

A recursive exponential average can also be used:

$$\bar{M}_k = \alpha M_k + (1-\alpha)\bar{M}_{k-1}.$$

If

$$\alpha$$

is small, the estimate is smooth but slow to respond.

If

$$\alpha$$

is large, the estimate responds rapidly but remains noisy.

Again,

$$\boxed{\text{stability} \leftrightarrow \text{responsiveness.}}$$

3.22 Multipath and the Meaning of the Best Beam

In a simple free-space environment, one might expect the best beam to point directly from the RN toward the BN.

Real environments often contain strong reflections.

For example,

$$\text{BN} \rightarrow \text{building} \rightarrow \text{RN}$$

may produce a stronger usable path than

$$\text{BN} \rightarrow \text{RN}.$$

This can occur if the direct path is obstructed.
Therefore the beam-search problem should generally be formulated as

$$\boxed{\text{find the beam producing the best link}}$$

rather than

$$\boxed{\text{find the geometric direction of the other radio.}}$$

This distinction is important in non-line-of-sight fixed wireless communication.

3.23 Multiple Useful Paths

A channel may contain several strong spatial paths.
For example,

$$\mathbf{h} = \alpha_1 \mathbf{a}(\theta_1) + \alpha_2 \mathbf{a}(\theta_2) + \alpha_3 \mathbf{a}(\theta_3).$$

Each term represents a different propagation path.
A beam search may therefore observe multiple local maxima.
One beam might correspond to the direct path while another corresponds to a strong reflection.
The strongest candidate may also depend on frequency.
Thus the spatial channel cannot always be represented by one simple direction.

3.24 False Peaks

Suppose the true beam-quality curve is

$$M(\theta).$$

The receiver does not observe this function perfectly.
Instead,

$$\hat{M}(\theta) = M(\theta) + \epsilon(\theta),$$

where

$$\epsilon(\theta)$$

represents measurement error.
Noise or interference may therefore cause a suboptimal beam to temporarily appear strongest.
If the system simply chooses

$$\arg\max_{\theta} \hat{M}(\theta)$$

from one short observation, the selected beam may be unstable.
Practical systems often reduce this problem using

- repeated measurements,
- averaging,

- candidate verification,
- hysteresis,
- confidence thresholds.

3.25 Hysteresis

Suppose the currently selected beam has metric

$$M_{\text{current}}$$

and another beam has metric

$$M_{\text{new}}.$$

Without hysteresis, the receiver may switch whenever

$$M_{\text{new}} > M_{\text{current}}.$$

If the two beams have nearly equal quality, measurement noise may cause frequent switching.
Instead, the system can require

$$M_{\text{new}} > M_{\text{current}} + \Delta,$$

where

$$\Delta$$

is a switching margin.
This reduces unnecessary beam changes.
Thus,

$$\boxed{\text{hysteresis trades rapid adaptation for stability.}}$$

3.26 Candidate Verification

Another strategy is to separate discovery from final selection.

The first stage may identify the best few beams:

$$\mathcal{C} = \{\mathbf{w}_{c_1}, \mathbf{w}_{c_2}, \ldots\}.$$

The receiver then spends additional time measuring only those candidates.

This avoids spending long measurement intervals on obviously poor beams while reducing the probability of making a decision based on a single noisy measurement.

Conceptually,

$$\boxed{\text{fast screening} \rightarrow \text{shortlist} \rightarrow \text{accurate verification.}}$$

3.27 Acquisition Versus Tracking

Acquisition and tracking are related but distinct problems.

During initial acquisition, the system may have little prior information.

It may need to search a large portion of the beam codebook.

After a link has been established, the current beam provides a strong prior.

Suppose the current beam index is

$$k.$$

Instead of searching all beams, the system may test

$$k-1, \qquad k, \qquad k+1.$$

This is a local search.

Thus,

$$\boxed{\text{acquisition} = \text{global search}}$$

while

$$\boxed{\text{tracking} = \text{local refinement.}}$$

Local tracking is generally much faster.

3.28 Why Beam Tracking Is Still Needed in Fixed Wireless

Even when the RN and BN do not physically move, the best beam may change.

Environmental changes can alter the channel:

- trees move,
- vehicles appear and disappear,
- construction changes reflection paths,
- weather changes propagation,
- interference sources change.

The effective channel is therefore

$$\mathbf{h}(t)$$

rather than a perfectly constant

$$\mathbf{h}.$$

A practical system periodically verifies that the current beam remains suitable.

3.29 Beam Search and Receive Gain

Beam acquisition interacts strongly with receiver gain control.

Different beam configurations may produce very different received amplitudes.

Suppose one beam has poor alignment:

$$P_{\mathrm{RX}} = -90 \text{ dBm}.$$

Another beam may align strongly:

$$P_{\mathrm{RX}} = -50 \text{ dBm}.$$

The difference is

$$40 \text{ dB}.$$

A receive gain suitable for the weak beam may cause the strong beam to saturate the receiver.

Conversely, a gain setting safe for the strong beam may place the weak beam close to the receiver noise floor.

Thus,

$$\boxed{\text{beam search} \leftrightarrow \text{gain control}.}$$

These cannot always be designed independently.

3.30 Gain Normalization During Beam Comparison

Suppose Beam A is measured with analog gain

$$G_A$$

and Beam B with gain

$$G_B.$$

Their raw digital amplitudes cannot be directly compared if

$$G_A \neq G_B.$$

In logarithmic units, one may conceptually reconstruct input power using

$$P_{\mathrm{input}} \approx P_{\mathrm{digital}} - G_{\mathrm{RX}}.$$

The receiver must therefore remember the gain state associated with each measurement.

Otherwise,

$$\boxed{\text{beam quality}}$$

becomes confused with

$$\boxed{\text{receiver amplification}.}$$

This is particularly important when the gain-control loop changes state during a beam sweep.

3.31 Beam-Switching and Settling Time

Changing a beam may not be instantaneous.

Hardware may require time for

- phase shifters to settle,
- amplifiers to settle,
- AGC to respond,
- filters to stabilize,
- synchronization estimates to converge.

Therefore a single beam measurement may consist of

$$\text{configure beam} \rightarrow \text{wait} \rightarrow \text{measure}.$$

If beam-switching time is

$$T_b,$$

gain-settling time is

$$T_g,$$

and actual measurement time is

$$T_m,$$

then one candidate may require

$$T_{\text{candidate}} = T_b + T_g + T_m.$$

An exhaustive search then requires approximately

$$\boxed{T_{\text{search}} = N_t N_r \left(T_b + T_g + T_m\right).}$$

This equation shows that hardware dynamics can be just as important as algorithmic complexity.

3.32 Detection Thresholds

During discovery, the receiver must decide whether a candidate beam contains a real signal.

Suppose the detection statistic is

$$M.$$

A simple detector may declare

$$\text{signal present}$$

if

$$M > \gamma,$$

where

$$\gamma$$

is a threshold.
Setting this threshold introduces another classic tradeoff.
If

$$\gamma$$

is too low, the receiver produces many false detections.
If

$$\gamma$$

is too high, weak real signals are missed.
These outcomes are described by

$$P_{\mathrm{FA}} = \text{probability of false alarm}$$

and

$$P_{\mathrm{D}} = \text{probability of detection.}$$

A practical acquisition system must choose a threshold that appropriately balances these probabilities.

3.33 The Acquisition Problem as Hypothesis Testing

Beam discovery can therefore be interpreted statistically.
For a particular beam, consider two hypotheses:

$$\mathcal{H}_0 : \text{desired signal absent,}$$

and

$$\mathcal{H}_1 : \text{desired signal present.}$$

The receiver observes some statistic

$$M$$

and decides between the two hypotheses.
For example,

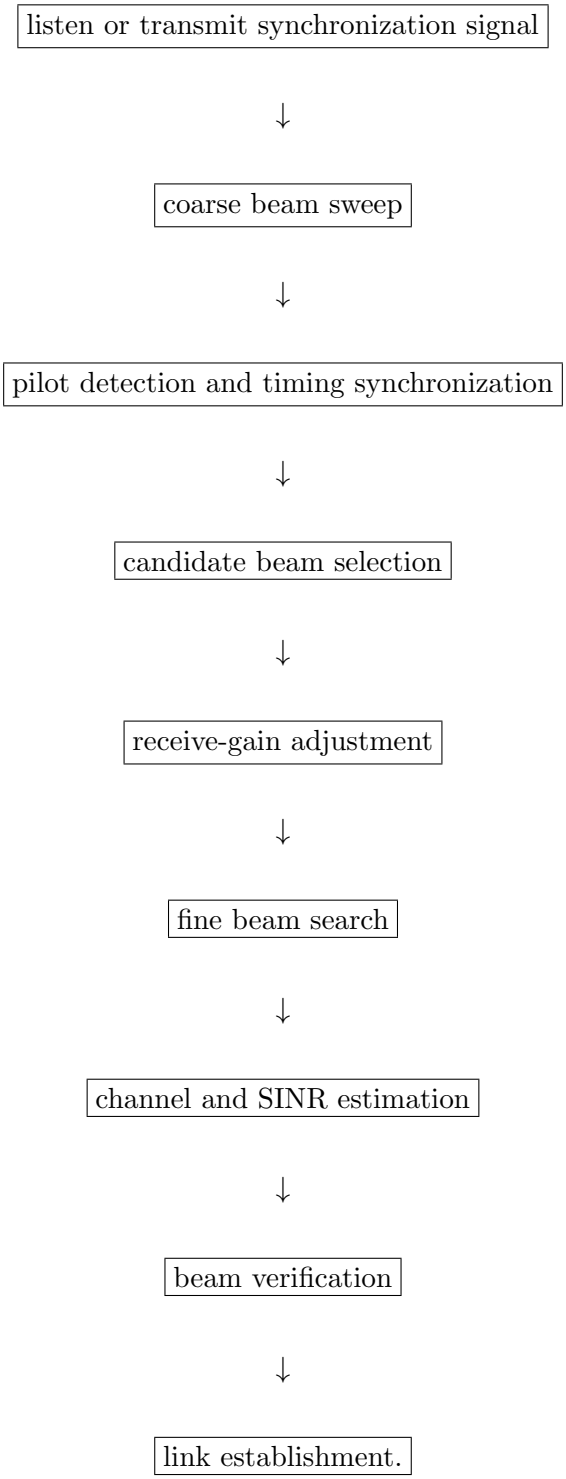
$$M = |C(\tau)|^2$$

could be a pilot-correlation energy.
The acquisition problem is therefore not simply geometrical.
It is also a statistical detection problem performed under uncertainty.

3.34 The Complete Acquisition Sequence

A practical acquisition process may contain several stages.

A simplified sequence is



Once communication begins, the system transitions from acquisition to tracking.

3.35 The Central Acquisition Tradeoffs

Beam acquisition is governed by several competing objectives.

3.35.1 Coverage Versus Gain

$$\boxed{\text{wide beam} \leftrightarrow \text{narrow beam}}$$

Wide beams simplify discovery but provide less gain.
Narrow beams provide greater gain but require more search.

3.35.2 Search Completeness Versus Speed

$$\boxed{\text{exhaustive search} \leftrightarrow \text{hierarchical search}}$$

Exhaustive search is robust but expensive.
Hierarchical search is faster but may miss the best candidate.

3.35.3 Measurement Accuracy Versus Latency

$$\boxed{\text{long dwell} \leftrightarrow \text{short dwell}}$$

Long measurements improve confidence but increase acquisition time.

3.35.4 Responsiveness Versus Stability

$$\boxed{\text{rapid switching} \leftrightarrow \text{hysteresis and averaging}}$$

Rapid adaptation follows channel changes quickly but may react to noise.
Strong averaging produces stable decisions but responds slowly.

3.35.5 Signal Gain Versus Interference Rejection

$$\boxed{\text{maximum desired power} \neq \text{maximum SINR.}}$$

The strongest signal path is not necessarily the path providing the best communication performance.

3.36 A Useful Mathematical View

Suppose each transmit and receive beam pair produces an effective SINR

$$\gamma_{ij}.$$

The ideal codebook search would select

$$\boxed{(i^*, j^*) = \arg\max_{i,j} \gamma_{ij}.}$$

In reality, the receiver does not know

$$\gamma_{ij}.$$

It observes noisy estimates

$$\hat{\gamma}_{ij}.$$

Furthermore, obtaining each estimate costs time.

The true engineering problem is therefore closer to

$$\boxed{\text{find a high-quality beam pair}}$$

subject to constraints on

$$\boxed{\text{measurement time,} \quad \text{hardware settling,} \quad \text{noise,} \quad \text{interference,} \quad \text{channel variation.}}$$

This makes beam acquisition an optimization problem under uncertainty.

3.37 Fundamental Perspective

The most important idea in beam acquisition is that the radio is trying to learn the spatial structure of an initially unknown communication channel.

The antenna array provides spatial degrees of freedom.

The codebook provides candidate spatial filters.

Measurements provide evidence about which filters reveal the desired transmitter most clearly.

The basic process is

$$\boxed{\text{unknown spatial channel} \rightarrow \text{probe with beams} \rightarrow \text{measure responses} \rightarrow \text{estimate useful direction} \rightarrow \text{refine beam.}}$$

This is analogous to many estimation problems in engineering.

The system does not know the answer directly.

Instead, it probes the environment, collects measurements, and progressively reduces uncertainty.

For an RN–BN link, the ultimate objective is not merely to point the antennas toward one another.

It is to discover a spatial configuration that supports reliable communication in the presence of

- path loss,
- multipath,
- interference,
- noise,
- receiver dynamic-range limitations,
- synchronization uncertainty.

Once such a beam has been found, a second practical problem immediately appears.

The received amplitude may change dramatically between poor and well-aligned beam states.

The receiver must therefore control its gain carefully so that weak signals remain measurable while strong signals do not saturate the analog front end or ADC.

This leads naturally to the next chapter:

$$\boxed{\text{Receiver Gain, Dynamic Range, and Clipping.}}$$

Chapter 4

Receiver Gain, Dynamic Range, and Clipping

A wireless receiver must operate over a very large range of possible input signal strengths.

A weak signal may arrive only slightly above the receiver noise floor, while a nearby transmitter or a highly aligned beam may produce a signal many orders of magnitude stronger.

The receiver must therefore amplify weak signals enough to measure them accurately, while avoiding excessive amplification of strong signals.

This creates one of the central practical problems in radio design:

$$\boxed{\text{How much gain should the receiver apply?}}$$

If the gain is too small, the signal uses only a small portion of the ADC range and becomes more sensitive to quantization and receiver noise.

If the gain is too large, one or more stages of the receiver may become nonlinear or saturate.

A useful high-level view is

$$\boxed{\text{too little gain} \rightarrow \text{poor sensitivity}}$$

and

$$\boxed{\text{too much gain} \rightarrow \text{clipping and distortion.}}$$

The purpose of gain control is to keep the received signal inside a useful operating region between these two extremes.

4.1 The Receiver Signal Chain

Before discussing gain control, it is useful to understand where gain is applied.

A simplified radio receiver may be represented as

$$\text{antenna} \rightarrow \text{LNA} \rightarrow \text{mixer} \rightarrow \text{filter} \rightarrow \text{VGA} \rightarrow \text{ADC} \rightarrow \text{digital processing.}$$

Each component performs a different function.

The antenna converts an electromagnetic wave into an electrical signal.

The Low-Noise Amplifier, or LNA, increases the level of very weak received signals while attempting to add as little noise as possible.

The mixer converts the signal from its RF carrier frequency to a lower intermediate frequency or directly to complex baseband.

Filters remove unwanted frequency components.

A Variable Gain Amplifier, or VGA, adjusts the signal amplitude before analog-to-digital conversion.

Finally, the ADC converts the continuous analog waveform into digital samples.

The overall analog receive gain can be represented as

$$G_{\mathrm{RX}}.$$

A simplified received signal at the ADC input is

$$y[n] = G_{\mathrm{RX}} \left(hx[n] + i[n] + n[n] \right),$$

where

- $x[n]$ is the desired transmitted signal,
- h is the desired channel,
- $i[n]$ represents interference,
- $n[n]$ represents receiver noise,
- G_{RX} is the total receive gain.

The gain multiplies everything entering the receiver.

This is important.

The receiver cannot generally amplify the desired signal without also amplifying interference and noise.

4.2 Why Gain Is Necessary

Signals received by an antenna can be extremely small.

Suppose the desired received signal has amplitude

$$A_{\mathrm{RF}}.$$

If the ADC has a full-scale voltage

$$V_{\mathrm{FS}},$$

and

$$A_{\mathrm{RF}} \ll V_{\mathrm{FS}},$$

then the raw signal may occupy only a very small fraction of the ADC range.

For example, suppose an ADC accepts values between approximately

$$-V_{\mathrm{FS}}$$

and

$$+V_{\mathrm{FS}}.$$

If the received signal uses only one percent of that range, then most of the ADC's representable levels are not being used.

Amplification attempts to scale the signal into a more useful region:

$$A_{\mathrm{ADC}} = G_{\mathrm{RX}} A_{\mathrm{RF}}.$$

Ideally,

$$A_{\mathrm{ADC}}$$

should be large enough to make efficient use of the converter while still leaving sufficient room for unpredictable peaks.

4.3 Dynamic Range

The term *dynamic range* describes the range between the smallest and largest signal levels that a system can handle effectively.

At the low end, signals are limited by noise and quantization.

At the high end, signals are limited by saturation and nonlinearity.

Conceptually,

$$\boxed{\text{noise floor} < \text{useful operating region} < \text{saturation level.}}$$

The dynamic range can be expressed as a ratio between the largest and smallest usable signal levels.

If the maximum usable power is $P_{\max}$ and the minimum usable power is $P_{\min}$, then

$$\mathrm{DR} = \frac{P_{\max}}{P_{\min}}.$$

In decibels,

$$\mathrm{DR}_{\mathrm{dB}} = 10 \log_{10} \left(\frac{P_{\max}}{P_{\min}} \right).$$

For voltage quantities measured across the same impedance,

$$\mathrm{DR}_{\mathrm{dB}} = 20 \log_{10} \left(\frac{V_{\max}}{V_{\min}} \right).$$

Wireless receivers often require large dynamic range because received signals can vary substantially due to

- distance,
- path loss,
- fading,
- beam direction,
- interference,
- transmit power differences.

4.4 The ADC Full-Scale Limit

An ADC can represent only a finite range of analog input voltages.

Suppose its allowed range is

$$-V_{\mathrm{FS}} \le y[n] \le V_{\mathrm{FS}}.$$

As long as

$$|y[n]| < V_{\mathrm{FS}},$$

the ADC can represent the input normally.

If instead

$$|y[n]| > V_{\mathrm{FS}},$$

the input exceeds the converter range.

The output saturates.

A simple clipping model is

$$y_{\mathrm{clip}}[n] = \begin{cases} V_{\mathrm{FS}}, & y[n] > V_{\mathrm{FS}}, \\ y[n], & |y[n]| \le V_{\mathrm{FS}}, \\ -V_{\mathrm{FS}}, & y[n] < -V_{\mathrm{FS}}. \end{cases}$$

This operation destroys information about the original waveform.

For example, inputs of

$$1.1V_{\mathrm{FS}}$$

and

$$2V_{\mathrm{FS}}$$

may both produce exactly

$$V_{\mathrm{FS}}.$$

The receiver can no longer determine what the original amplitudes were.

4.5 Why Clipping Is Especially Harmful

Clipping is not simply a small amplitude error.

It is a nonlinear operation.

Consider an ideal linear system:

$$y = ax.$$

If the input doubles, the output doubles.

For a clipping system,

$$y = \mathrm{clip}(ax),$$

this relationship eventually fails.
The shape of the waveform changes.
This distortion can affect

- amplitude estimation,
- phase estimation,
- channel estimation,
- synchronization,
- EVM,
- symbol detection,
- SINR estimation.

For modulated communication signals, clipping can therefore produce errors across many transmitted symbols.

4.6 Clipping and Spectral Distortion

Nonlinear operations generate new frequency components.
Suppose a sinusoid

$$x(t) = A\cos(2\pi f t)$$

is clipped.
The resulting waveform is no longer a pure sinusoid.
It contains harmonics at frequencies such as

$$2f, \qquad 3f, \qquad 4f,$$

and others depending on the exact nonlinearity.
For a broadband communication signal, the distortion can spread energy across the spectrum.
This is especially important for OFDM, where many subcarriers exist close together.
Clipping in the time domain can therefore create interference across multiple subcarriers.
Thus,

$$\boxed{\text{clipping} \rightarrow \text{nonlinear distortion} \rightarrow \text{spectral spreading} \rightarrow \text{worse demodulation.}}$$

4.7 Quantization

The ADC does not represent every possible analog voltage exactly.
Instead, it maps the continuous input voltage to one of a finite number of digital levels.
For a B-bit ADC, the number of possible levels is

$$2^B.$$

For example,

$$B = 8$$

provides

$$2^8 = 256$$

levels.

A 12-bit converter provides

$$2^{12} = 4096$$

levels.

The difference between the true input and the selected ADC level is called the quantization error.

A simplified model is

$$y_q[n] = y[n] + q[n],$$

where

$$q[n]$$

represents quantization error.

4.8 Quantization Step Size

Suppose the total ADC voltage range is

$$2V_{\text{FS}}.$$

For an ideal B-bit converter, the approximate quantization step is

$$\Delta = \frac{2V_{\text{FS}}}{2^B}.$$

As the number of bits increases,

$$B \uparrow \quad \Rightarrow \quad \Delta \downarrow .$$

Thus higher-resolution ADCs can represent smaller changes in input voltage.

However, ADC resolution alone does not solve the gain problem.

Even a high-resolution ADC performs poorly if the received signal uses only a tiny fraction of its full-scale range.

4.9 Effective Use of ADC Range

Suppose the ADC has full-scale amplitude

$$V_{\text{FS}}.$$

If the signal peak is

$$V_{\text{signal}} = V_{\text{FS}},$$

the available converter range is used efficiently.

Now suppose instead

$$V_{\text{signal}} = \frac{V_{\text{FS}}}{16}.$$

A factor of 16 corresponds to

$$2^4.$$

The signal therefore uses roughly four fewer bits of amplitude range.

In this simplified sense, under-driving an ADC can reduce the effective number of useful bits.

This is one reason why gain control attempts to keep the input reasonably close to full scale without actually clipping.

4.10 Quantization SNR

For an ideal B-bit ADC driven by a full-scale sinusoid, a commonly used approximation is

$$\boxed{\text{SNR}_q \approx 6.02B + 1.76 \quad \text{dB}.}$$

Each additional ADC bit therefore improves ideal quantization SNR by approximately

$$6 \text{ dB}.$$

For example, increasing the converter resolution from 10 to 11 bits ideally improves quantization SNR by approximately

$$6 \text{ dB}.$$

However, this formula assumes effective use of the full ADC range.

If the signal level is much lower, the practical quantization SNR is reduced.

4.11 The Fundamental Gain-Control Tradeoff

The receiver therefore faces two opposing constraints.

Increasing gain improves ADC utilization:

$$G_{\text{RX}} \uparrow \rightarrow \text{larger ADC signal}.$$

But increasing gain also moves the system closer to saturation:

$$G_{\text{RX}} \uparrow \rightarrow \text{less headroom}.$$

Thus,

$$\boxed{\text{too little gain} \rightarrow \text{poor ADC utilization}}$$

while

$$\boxed{\text{too much gain} \rightarrow \text{clipping.}}$$

The goal is to choose

$$G_{\mathrm{RX}}$$

such that the signal occupies a substantial fraction of the ADC range while maintaining enough margin for unexpected peaks.

4.12 Average Power and Peak Power

An important distinction is the difference between average signal power and instantaneous peak power.

Suppose the average power is

$$P_{\mathrm{avg}}.$$

The waveform may occasionally reach a much larger instantaneous value

$$P_{\mathrm{peak}}.$$

The ratio is the Peak-to-Average Power Ratio:

$$\mathrm{PAPR} = \frac{P_{\mathrm{peak}}}{P_{\mathrm{avg}}}.$$

In decibels,

$$\mathrm{PAPR}_{\mathrm{dB}} = 10 \log_{10} \left(\frac{P_{\mathrm{peak}}}{P_{\mathrm{avg}}} \right).$$

This means that setting the average signal close to ADC full scale can be dangerous.
Even if

$$P_{\mathrm{avg}} < P_{\mathrm{FS}},$$

individual waveform peaks may still satisfy

$$P_{\mathrm{peak}} > P_{\mathrm{FS}}.$$

The receiver therefore requires headroom.

4.13 Headroom

Suppose the ADC full-scale power is

$$P_{\mathrm{FS}}.$$

The receiver may choose a target average level

$$\boxed{P_{\mathrm{target}} = P_{\mathrm{FS}} - H,}$$

where H is the desired headroom in decibels.
For example, if

$$P_{\mathrm{FS}} = 0 \text{ dBFS}$$

and the chosen headroom is

$$H = 10 \text{ dB},$$

then the target average level is

$$P_{\mathrm{target}} = -10 \text{ dBFS}.$$

This margin provides room for

- waveform peaks,
- fading,
- sudden beam-alignment gain,
- transient interference,
- AGC response delay,
- estimation uncertainty.

The correct headroom is therefore a system-design choice.
Too much headroom wastes ADC range.
Too little headroom increases clipping probability.

4.14 Automatic Gain Control

Because received power can vary over time, the receiver usually cannot rely on one fixed gain value.
Instead, it uses Automatic Gain Control, or AGC.
The basic AGC loop is

$$\boxed{\text{measure signal level} \rightarrow \text{compare with target} \rightarrow \text{adjust gain} \rightarrow \text{measure again.}}$$

Suppose the measured receive power at time step k is

$$P_k.$$

The desired target is

$$P_{\mathrm{target}}.$$

Define the error

$$e_k = P_{\mathrm{target}} - P_k.$$

A simple gain-control law is

$$G_{k+1} = G_k + \mu e_k,$$

where

$$\mu$$

controls how aggressively the gain changes.
If the signal is too weak,

$$P_k < P_{\text{target}},$$

then

$$e_k > 0$$

and the receiver increases gain.
If the signal is too strong,

$$P_k > P_{\text{target}},$$

then

$$e_k < 0$$

and the receiver reduces gain.
Thus the AGC behaves as a feedback control system.

4.15 AGC as a Feedback Loop

The AGC can be viewed using the same concepts as a general control system.
The plant is the receiver gain chain.
The measured output is signal level.
The reference is the target level.
The controller adjusts the gain.
Conceptually,

$$\boxed{P_{\text{target}} \rightarrow \text{controller} \rightarrow G_{\text{RX}} \rightarrow P_{\text{measured}} \rightarrow \text{feedback.}}$$

The controller must balance speed and stability.
If gain changes too slowly, the receiver may remain saturated for too long.
If gain changes too aggressively, the gain can oscillate around the desired value.

4.16 Fast Attack and Slow Release

The cost of excessive gain is usually asymmetric.
If the gain is too high, the ADC may clip immediately.
If the gain is slightly too low, the signal remains measurable but with reduced resolution.
This motivates the common strategy known as

$$\boxed{\text{fast attack, slow release.}}$$

When the signal suddenly becomes stronger,

$$G_{\mathrm{RX}} \downarrow$$

quickly.
When the signal becomes weaker,

$$G_{\mathrm{RX}} \uparrow$$

more gradually.
A simple model is

$$G_{k+1} = G_k + \mu_k \left(P_{\mathrm{target}} - P_k \right),$$

where

$$\mu_k = \begin{cases} \mu_{\mathrm{attack}}, & P_k > P_{\mathrm{target}}, \\ \mu_{\mathrm{release}}, & P_k \leq P_{\mathrm{target}}. \end{cases}$$

Typically,

$$\boxed{\mu_{\mathrm{attack}} > \mu_{\mathrm{release}}.}$$

The receiver therefore reacts aggressively to possible saturation but cautiously when increasing amplification.

4.17 Why Beam Search Makes AGC Difficult

Beam acquisition can create very large changes in received signal level.
Suppose one beam is poorly aligned.
The received signal may be

$$-90 \text{ dBm}.$$

The next beam may align with a strong propagation path and produce

$$-50 \text{ dBm}.$$

The difference is

$$40 \text{ dB}.$$

In linear power terms,

$$10^{40/10} = 10^4.$$

Thus the received power has increased by a factor of

$$10,000.$$

A gain setting appropriate for the first beam may be completely inappropriate for the second.
This explains why beam acquisition and gain control are tightly coupled.

4.18 Gain State and Signal Measurement

Suppose Beam A is measured using gain

$$G_A$$

while Beam B is measured using

$$G_B.$$

The resulting digital powers might be

$$P_{A,\text{digital}}$$

and

$$P_{B,\text{digital}}.$$

These values cannot be compared directly if

$$G_A \neq G_B.$$

In logarithmic units, the input-referred received power can be approximated as

$$\boxed{P_{\text{RF}} \approx P_{\text{digital}} - G_{\text{RX}}.}$$

For example, suppose Beam A produces

$$P_{A,\text{digital}} = -10 \text{ dBFS}$$

using

$$G_A = 50 \text{ dB}.$$

Beam B also produces

$$P_{B,\text{digital}} = -10 \text{ dBFS},$$

but uses only

$$G_B = 30 \text{ dB}.$$

Their digital amplitudes are identical, but Beam B required

$$20 \text{ dB}$$

less receiver amplification.
Beam B therefore corresponds to a much stronger RF input.
The gain state must therefore be included in beam-scoring calculations.

4.19 Analog Gain and Digital Gain

Gain may be applied in both analog and digital domains.

Analog gain occurs before the ADC:

$$x_{\text{analog}} \to G_{\text{analog}} \to \text{ADC}.$$

Digital gain occurs after conversion:

$$x_{\text{digital}} \to G_{\text{digital}} x_{\text{digital}}.$$

These two types of gain are not equivalent.

Analog gain increases the signal amplitude before quantization and can therefore improve ADC utilization.

Digital gain cannot recover information lost because the signal was too small during conversion.

If the ADC recorded

$$x_q[n],$$

digital amplification produces

$$G_{\text{digital}} x_q[n].$$

It amplifies both the desired samples and any quantization error already present.

Similarly, digital gain cannot repair clipping that occurred before or inside the ADC.

Thus,

digital gain cannot restore information already lost in analog conversion.

4.20 Gain Distribution Across the Receiver

The total gain is often distributed across multiple components.

For example,

$$G_{\text{RX}} = G_{\text{LNA}} G_{\text{VGA}} G_{\text{other}}$$

in linear units.

In decibels,

$$G_{\text{RX,dB}} = G_{\text{LNA,dB}} + G_{\text{VGA,dB}} + G_{\text{other,dB}}.$$

The placement of gain matters.

Gain early in the receiver chain helps overcome noise introduced by later stages.

However, excessive early gain may cause later stages to saturate.

Receiver design therefore requires careful allocation of gain across the entire chain.

4.21 Noise Figure and Early Gain

A major reason for using a low-noise amplifier near the antenna is to preserve weak-signal sensitivity.

Suppose the first amplifier introduces relatively little additional noise while providing substantial gain.

The signal entering later stages is then much larger relative to the noise introduced by those stages.

This is why the noise performance of the first few receiver components is particularly important.

The complete analysis is often performed using the receiver noise figure and Friis noise equation, but the fundamental idea is

low-noise gain early in the chain helps preserve SNR.

However, this benefit must be balanced against linearity and saturation.

4.22 The ADC Is Not the Only Saturation Point

It is tempting to think that clipping occurs only at the ADC.

In reality, every analog component has a finite linear operating region.

For example,

$$\text{LNA} \rightarrow \text{mixer} \rightarrow \text{VGA} \rightarrow \text{ADC}$$

can each saturate or become nonlinear.

If the LNA becomes nonlinear, reducing gain later in the chain does not fix the distortion.

Likewise, if the mixer has already generated nonlinear products, the ADC faithfully digitizes an already-corrupted waveform.

Therefore,

receiver linearity must be maintained throughout the entire analog chain.

4.23 Compression

An ideal amplifier obeys

$$P_{\text{out}} = GP_{\text{in}}$$

over all input powers.

A real amplifier behaves this way only over a limited range.

At high input power, the output begins to increase more slowly than expected.

This behavior is called compression.

A common specification is the 1-dB compression point,

$$P_{\text{1dB}}.$$

At this point, the amplifier output is approximately

$$1 \text{ dB}$$

lower than what an ideal linear extrapolation predicts.

Conceptually,

$$\boxed{P_{1\text{dB}} = \text{approximate beginning of significant gain compression.}}$$

Operation well below this point is generally preferred when high linearity is required.

4.24 Intermodulation

Nonlinear receiver components can also cause different input frequencies to mix.

Suppose two strong signals exist at frequencies

$$f_1$$

and

$$f_2.$$

A nonlinear amplifier can generate new components such as

$$2f_1 - f_2$$

and

$$2f_2 - f_1.$$

These are third-order intermodulation products.

They are troublesome because they may fall inside the desired signal band.

The receiver may therefore observe interference that was not present at the antenna as an independent transmitter, but was created internally by analog nonlinearity.

4.25 The Third-Order Intercept Point

A common measure of nonlinear performance is the third-order intercept point.

When referred to the input, it is called

$$\text{IIP3}.$$

A larger IIP3 generally indicates that the receiver can tolerate stronger interfering signals before significant third-order distortion occurs.

The IIP3 is not normally an actual operating point.

Instead, it is obtained by extrapolating the linear growth of the fundamental signal and the faster growth of third-order distortion products.

The essential interpretation is

$$\boxed{\text{higher IIP3} \rightarrow \text{better large-signal linearity.}}$$

4.26 Interference-Induced Saturation

A difficult receiver condition occurs when the desired signal is weak but an interferer is very strong:

$$P_{\text{desired}} \ll P_{\text{interferer}}.$$

The desired signal may need substantial gain to become measurable.
However, that same gain is also applied to the interferer.
The strong interferer may then cause saturation.
For example,

$$\text{desired signal} = -90 \text{ dBm}$$

and

$$\text{interferer} = -40 \text{ dBm}$$

represent a difference of

$$50 \text{ dB}.$$

The interferer has

$$10^{50/10} = 100,000$$

times more power.

The receiver cannot simply increase gain until the desired signal becomes large because the interferer may saturate the front end first.

4.27 Why Digital Interference Cancellation Has Limits

Suppose the analog receiver remains linear.
The ADC observes

$$y = s + i + n.$$

A digital algorithm may then estimate

$$i$$

and attempt to subtract it.
However, suppose the interferer causes analog clipping before digitization.
Then the ADC observes something closer to

$$y_{\text{ADC}} = \text{clip}(s + i + n).$$

The clipping operation is nonlinear.
In general,

$$\text{clip}(s + i + n) - i \neq s + n.$$

Information has already been destroyed.

Thus,

$$\boxed{\text{digital cancellation cannot recover information lost through analog saturation.}}$$

This is why analog dynamic range remains important even in systems with sophisticated digital interference cancellation.

4.28 Gain Settling

Changing receiver gain is not necessarily instantaneous.

Analog components require some finite time to settle to a new state.

Suppose the receiver changes gain at time

$$t_0.$$

The output may contain a transient before converging to the expected value.

Measurements taken during this transient can be inaccurate.

Therefore a measurement sequence may be

$$\boxed{\text{set gain} \rightarrow \text{wait} \rightarrow \text{measure.}}$$

If the gain settling time is

$$T_{\text{settle}},$$

then measurements should generally avoid the unstable portion of the response.

4.29 Gain Settling During Beam Search

Beam search may require gain changes for many beam candidates.

Suppose each beam requires

- beam configuration time T_{beam},
- gain-update time T_{gain},
- settling time T_{settle},
- measurement time T_{measure}.

Then the total time per candidate is approximately

$$T_{\text{candidate}} = T_{\text{beam}} + T_{\text{gain}} + T_{\text{settle}} + T_{\text{measure}}.$$

For N_{beams} candidates,

$$\boxed{T_{\text{scan}} = N_{\text{beams}} \left(T_{\text{beam}} + T_{\text{gain}} + T_{\text{settle}} + T_{\text{measure}} \right).}$$

This shows that acquisition speed depends not only on DSP computation but also on analog hardware response time.

4.30 Measurement Windows

Gain control itself requires an estimate of signal level.

Suppose samples are

$$y[0], y[1], \ldots, y[N-1].$$

An average-power estimate is

$$\hat{P} = \frac{1}{N} \sum_{n=0}^{N-1} |y[n]|^2.$$

A larger N generally produces a more stable estimate.

However,

$$N \uparrow$$

also increases measurement delay.

Thus AGC contains the familiar tradeoff

$$\boxed{\text{accurate power estimate} \leftrightarrow \text{fast response.}}$$

A very short averaging window may react quickly to individual waveform peaks.

A very long averaging window may fail to respond quickly enough to a sudden strong signal.

4.31 Peak Detection Versus Average-Power Detection

An AGC may use different measures of signal amplitude.

One possibility is average power:

$$P_{\text{avg}} = \frac{1}{N} \sum_{n} |y[n]|^2.$$

Another is peak amplitude:

$$A_{\text{peak}} = \max_{n} |y[n]|.$$

Average power is useful for maintaining a stable operating level.

Peak detection is useful for identifying imminent clipping.

A practical receiver may therefore use both.

Conceptually,

$$\boxed{\text{average detector} \rightarrow \text{normal gain control}}$$

and

$$\boxed{\text{peak detector} \rightarrow \text{clipping protection.}}$$

4.32 Clipping Probability

For waveforms with varying amplitude, it may be impossible to guarantee that no sample ever exceeds full scale without using very large headroom.

Instead, a system may tolerate a very small clipping probability.

Define

$$P_{\text{clip}} = P\left(|y| > V_{\text{FS}}\right).$$

Increasing headroom reduces

$$P_{\text{clip}},$$

but also reduces average ADC utilization.

Thus gain design may be considered as a probabilistic tradeoff:

$$\boxed{\text{larger headroom} \rightarrow \text{less clipping} \rightarrow \text{poorer ADC utilization.}}$$

The desired operating point depends on the signal distribution and system requirements.

4.33 Gain Steps

Real gain hardware often cannot choose an arbitrary continuous value.

Instead, the receiver may support discrete gain settings:

$$G_1, G_2, G_3, \ldots, G_M.$$

For example,

$$0 \text{ dB}, \quad 5 \text{ dB}, \quad 10 \text{ dB}, \quad 15 \text{ dB}, \ldots$$

The AGC therefore performs a quantized control action.

If the gain step is large, the system converges quickly but may overshoot the ideal target.

If the gain step is small, control is more precise but may require more adjustments.

Thus,

$$\boxed{\text{large gain steps} \rightarrow \text{fast but coarse control}}$$

while

$$\boxed{\text{small gain steps} \rightarrow \text{slow but precise control.}}$$

4.34 Two-Stage Gain Control

A practical AGC may use multiple stages.

For example,

$$\text{coarse analog gain} + \text{fine digital scaling.}$$

The analog stage places the waveform into an appropriate ADC range.

The digital stage then adjusts scaling for later DSP blocks.

A conceptual sequence is

$$\boxed{\text{large input-range correction} \rightarrow \text{analog gain}}$$

followed by

$$\boxed{\text{fine normalization} \rightarrow \text{digital scaling.}}$$

This allows the receiver to combine large dynamic range with precise downstream signal scaling.

4.35 Fixed-Point Dynamic Range

Dynamic-range problems continue even after the ADC.

Many real-time DSP systems use fixed-point arithmetic.

Suppose a digital signal is represented by

$$B$$

bits.

There is again a largest representable magnitude.

If a calculation produces a value beyond this range, overflow occurs.

Beamforming is particularly relevant.

Suppose

$$N$$

antenna signals are added coherently.

In the worst case, amplitudes can add approximately as

$$A_{\text{out}} \approx N A_{\text{in}}.$$

The number of additional bits required to represent this increase is approximately

$$\boxed{\lceil \log_2 N \rceil}.$$

For example, if

$$N = 16,$$

then

$$\log_2(16) = 4.$$

Approximately four additional integer bits may therefore be required for the worst-case coherent sum.

Thus dynamic-range analysis applies to both

$$\boxed{\text{analog hardware}}$$

and

$$\boxed{\text{digital arithmetic.}}$$

4.36 Normalization

One way to control digital beamformer amplitude is to normalize the beamforming weights.
For example,

$$\mathbf{w}_{\text{norm}} = \frac{\mathbf{w}}{\|\mathbf{w}\|_2}.$$

This limits the overall magnitude of the weight vector.
However, the appropriate normalization depends on the physical constraint.
A system may wish to preserve

- total transmit power,
- per-antenna power,
- receive beamformer gain,
- digital headroom.

Therefore normalization is not simply a mathematical convenience.
It is part of managing signal magnitude throughout the processing chain.

4.37 Dynamic Range as a System-Level Problem

A useful mistake to avoid is thinking of dynamic range as an ADC-only problem.
A complete receiver must preserve signal integrity through

$$\boxed{\text{antenna} \to \text{LNA} \to \text{mixer} \to \text{filters} \to \text{VGA} \to \text{ADC} \to \text{digital processing.}}$$

Each stage has

- a noise floor,
- a maximum useful amplitude,
- finite gain,
- finite linearity.

The signal must remain in an acceptable operating region at every stage.
Thus the true design objective is not simply

$$\text{avoid ADC clipping.}$$

It is

$$\boxed{\text{preserve sufficient SNR while maintaining linearity throughout the receiver.}}$$

4.38 The Central Gain-Control Tradeoff

The gain-control problem can now be summarized.

Increasing gain produces

$$\text{better ADC utilization}$$

and can improve effective representation of weak signals.

But it also produces

$$\text{less headroom}$$

and increases the risk of

$$\text{saturation.}$$

Thus the receiver must continuously balance

$$\boxed{\text{sensitivity} \leftrightarrow \text{linearity.}}$$

This balance becomes particularly difficult when signal strength changes rapidly, as during beam acquisition.

4.39 Relationship to Beam Acquisition

The previous chapter described beam acquisition as a search over spatial configurations.

The present chapter shows that the measurements used during that search are affected by receiver gain.

The complete interaction is approximately

$$\boxed{\text{beam selection} \rightarrow \text{received power} \rightarrow \text{gain adjustment} \rightarrow \text{ADC signal level} \rightarrow \text{beam-quality measurement.}}$$

This creates a feedback relationship.

Changing the beam changes signal amplitude.

Changing the amplitude changes the appropriate gain.

Changing the gain changes the digital representation used to evaluate the beam.

Therefore gain state must be carefully tracked during beam acquisition.

4.40 Fundamental Perspective

Receiver gain control is fundamentally an attempt to map a very large analog input range into a much smaller usable electronic and digital range.

The outside world may produce signals spanning many tens of decibels.

The ADC and analog components have finite operating ranges.

The receiver therefore continuously rescales the signal.

Conceptually,

$$\boxed{\text{large RF dynamic range} \rightarrow \text{adaptive gain} \rightarrow \text{manageable ADC range.}}$$

The ideal receiver keeps the waveform

- large enough to measure accurately,
- small enough to remain linear,
- stable enough for channel estimation and demodulation.

The central principle is

amplify enough to preserve information, but never so much that information is destroyed.

Once the signal has been placed safely within the receiver's operating range, the next question is how to evaluate the quality of that signal.

This leads naturally to quantities such as

$$\text{RSSI}, \qquad \text{SNR}, \qquad \text{SINR},$$

which form the basis for the next stage of link-quality estimation.

Chapter 5

RSSI, SNR, and SINR

Wireless receivers need practical metrics for describing signal strength and signal quality.

Three of the most common are

$$\text{RSSI}, \qquad \text{SNR}, \qquad \text{SINR}.$$

Although these quantities are related, they answer different questions.

A receiver may observe a strong signal that is difficult to decode, or a relatively weak signal that is nevertheless very clean.

For this reason, understanding the difference between received power and usable signal quality is fundamental to practical radio design.

5.1 Received Signal Strength Indicator

RSSI stands for Received Signal Strength Indicator.

It is a measure related to the amount of power observed by the receiver.

A simplified model is

$$P_{\text{RSSI}} \approx P_S + P_I + P_N,$$

where

- P_S is desired-signal power,
- P_I is interference power,
- P_N is receiver and thermal noise power.

RSSI therefore answers approximately

> How much total energy is present at the receiver?

It does not necessarily answer

> How easy is it to recover the desired data?

This distinction is extremely important.

A large RSSI may be caused by

- a strong desired signal,
- a strong interferer,
- several nearby transmitters,
- broadband noise,
- some combination of these.

Therefore,

high RSSI does not necessarily imply a high-quality link.

5.2 RSSI in Practice

In a practical receiver, RSSI is often derived from a measurement of average received power.

If complex baseband samples are

$$y[n],$$

then an average digital power estimate may be

$$\hat{P}_y = \frac{1}{N} \sum_{n=0}^{N-1} |y[n]|^2.$$

If the receiver gain is known, this value can be converted into an estimate of input-referred RF power.

In logarithmic units, one may approximately write

$$P_{\mathrm{RF,dBm}} \approx P_{\mathrm{digital}} - G_{\mathrm{RX}} + C,$$

where

- G_{RX} is the analog receive gain,
- C represents calibration constants associated with the hardware.

This is why RSSI measurements often depend on receiver calibration.

A digital magnitude by itself is not sufficient to determine actual received RF power unless the gain state of the receiver is also known.

5.3 The Decibel Scale

Wireless powers are commonly expressed in decibels because received powers can span many orders of magnitude.

For a power ratio,

$$R = \frac{P_1}{P_2},$$

the corresponding decibel value is

$$R_{\text{dB}} = 10 \log_{10} \left(\frac{P_1}{P_2} \right).$$

Absolute power is often expressed in dBm, which is referenced to one milliwatt:

$$P_{\text{dBm}} = 10 \log_{10} \left(\frac{P}{1 \text{ mW}} \right).$$

For example,

$$1 \text{ mW} = 0 \text{ dBm},$$

$$0.1 \text{ mW} = -10 \text{ dBm},$$

and

$$0.001 \text{ mW} = -30 \text{ dBm}.$$

A difference of

$$10 \text{ dB}$$

corresponds to a factor of ten in power.
A difference of

$$20 \text{ dB}$$

corresponds to a factor of one hundred.
This logarithmic representation makes wireless link budgets much easier to work with.

5.4 Signal-to-Noise Ratio

Signal-to-Noise Ratio compares desired-signal power with noise power.

It is defined as

$$\boxed{\text{SNR} = \frac{P_S}{P_N}.}$$

In decibels,

$$\text{SNR}_{\text{dB}} = 10 \log_{10} \left(\frac{P_S}{P_N} \right).$$

Suppose the desired received signal has power

$$P_S = -60 \text{ dBm}$$

and the noise power is

$$P_N = -90 \text{ dBm}.$$

Then

$$\mathrm{SNR} \approx 30 \text{ dB}.$$

The subtraction is possible because both values are already expressed logarithmically:

$$\mathrm{SNR}_{\mathrm{dB}} = P_{S,\mathrm{dBm}} - P_{N,\mathrm{dBm}}.$$

Thus,

$$-60 - (-90) = 30 \text{ dB}.$$

5.5 What Determines the Noise Floor

A useful approximation for thermal noise power is

$$P_N = kTB,$$

where

- k is Boltzmann's constant,
- T is absolute temperature,
- B is receiver bandwidth.

At approximately room temperature, thermal noise density is often approximated as

$$-174 \text{ dBm/Hz}.$$

For bandwidth B,

$$P_{N,\mathrm{thermal}} \approx -174 + 10\log_{10}(B) \quad \text{dBm}.$$

For example, if

$$B = 20 \text{ MHz},$$

then

$$10\log_{10}(20 \times 10^6) \approx 73 \text{ dB}.$$

Therefore the ideal thermal noise power is approximately

$$-174 + 73 = -101 \text{ dBm}.$$

A real receiver adds its own noise.
If the receiver noise figure is

$$NF,$$

then the practical noise floor becomes approximately

$$\boxed{P_N \approx -174 + 10\log_{10}(B) + NF}$$

in dBm.
This equation is useful in practical link-budget calculations.

5.6 Why SNR Is Useful

SNR describes how clearly the desired signal stands above random noise.

As SNR increases, the received constellation points become more tightly clustered around their ideal positions.

At low SNR,

$$\text{noise magnitude} \sim \text{distance between constellation points},$$

and symbol errors become likely.

At high SNR,

$$\text{noise magnitude} \ll \text{distance between constellation points},$$

and the receiver can distinguish symbols much more reliably.

Thus SNR is closely related to

- bit error rate,
- symbol error rate,
- achievable modulation order,
- decoder performance.

However, SNR assumes that the main impairment is noise.

In many real wireless networks, this assumption is incomplete.

5.7 Interference

A receiver may observe other structured radio signals in addition to the desired transmission.

Suppose

$$y = h_s x_s + h_i x_i + n.$$

The desired component is

$$h_s x_s,$$

while

$$h_i x_i$$

is interference.

Unlike thermal noise, interference may itself be a valid communication waveform.

It may originate from

- another customer,
- another base station,
- another network,

- another technology sharing the same band,
- leakage from a nearby channel.

If interference is strong, an SNR calculation that considers only thermal noise can significantly overestimate actual link quality.

5.8 Signal-to-Interference-plus-Noise Ratio

The more useful quantity in an interference-limited environment is Signal-to-Interference-plus-Noise Ratio.

It is defined as

$$\boxed{\text{SINR} = \frac{P_S}{P_I + P_N}.}$$

In decibels,

$$\text{SINR}_{\text{dB}} = 10 \log_{10} \left(\frac{P_S}{P_I + P_N} \right).$$

SINR answers the practical question

> How strong is the desired signal compared with everything that makes it harder to decode?

The conceptual distinction is

$$\boxed{\text{RSSI: total received energy}}$$

$$\boxed{\text{SNR: desired signal versus noise}}$$

$$\boxed{\text{SINR: desired signal versus interference and noise.}}$$

5.9 Why RSSI and SINR Can Disagree

Consider two beams.

Beam	P_S	P_{I+N}
A	-50 dBm	-80 dBm
B	-45 dBm	-47 dBm

Beam B contains a stronger desired signal.
However, it also contains much stronger interference.
For Beam A,

$$\text{SINR}_A \approx -50 - (-80) = 30 \text{ dB}.$$

For Beam B,

$$\mathrm{SINR}_B \approx -45 - (-47) = 2 \text{ dB}.$$

Beam B may produce a larger RSSI, but Beam A provides a much cleaner communication channel.

This is one of the most important practical distinctions in wireless systems.

5.10 RSSI During Beam Search

RSSI is still useful during beam acquisition.

It is inexpensive to calculate and can rapidly eliminate very weak candidate beams.

For example, suppose a beam sweep produces

$$-92, \quad -88, \quad -61, \quad -59, \quad -90 \text{ dBm}.$$

The receiver can immediately conclude that the third and fourth beams are more promising than the others.

Thus RSSI can be useful for

$$\boxed{\text{coarse candidate selection.}}$$

However, choosing between the remaining candidates may require a more informative metric.

A practical search may therefore use

$$\text{RSSI} \rightarrow \text{candidate filtering}$$

followed by

$$\text{SINR or EVM} \rightarrow \text{final selection.}$$

5.11 Post-Beamforming SINR

For a multi-antenna receiver, SINR should generally be considered after spatial combining.

Suppose

$$\mathbf{y} = \mathbf{h}_s x_s + \mathbf{i} + \mathbf{n}.$$

The beamformer produces

$$z = \mathbf{w}^H \mathbf{y}.$$

The desired-signal power after beamforming is

$$P_S = P_x \left| \mathbf{w}^H \mathbf{h}_s \right|^2.$$

If the interference-plus-noise covariance is

$$\mathbf{R}_{I+N},$$

then the corresponding output power is

$$P_{I+N} = \mathbf{w}^H \mathbf{R}_{I+N} \mathbf{w}.$$

Therefore,

$$\boxed{\mathrm{SINR}(\mathbf{w}) = \frac{P_x \left|\mathbf{w}^H \mathbf{h}_s\right|^2}{\mathbf{w}^H \mathbf{R}_{I+N} \mathbf{w}}.}$$

This equation explains why beamforming should not generally optimize only desired-signal power.

Changing the beam changes both

$$P_S$$

and

$$P_I.$$

5.12 Beam Selection

If beam selection is based only on RSSI, the objective may be written as

$$\mathbf{w}^* = \arg\max_{\mathbf{w}} \mathrm{RSSI}(\mathbf{w}).$$

This may favor beams containing large interfering signals.

A more meaningful objective is often

$$\boxed{\mathbf{w}^* = \arg\max_{\mathbf{w}} \mathrm{SINR}(\mathbf{w}).}$$

This allows the receiver to trade a small amount of desired-signal gain for substantial interference rejection.

For example,

$$P_S \downarrow 2 \text{ dB}$$

may be worthwhile if

$$P_I \downarrow 15 \text{ dB}.$$

The resulting SINR improves significantly.

5.13 How SINR Is Estimated in Practice

The receiver usually cannot directly measure

$$P_S, \qquad P_I, \qquad P_N$$

as completely separate physical quantities.

It observes only the combined waveform.

Therefore SINR is normally estimated indirectly.

One common method uses known pilot symbols.
Suppose

$$y_k = h_k x_k + i_k + n_k,$$

where x_k is known.
The receiver estimates the channel:

$$\hat{h}_k.$$

The expected desired signal is then

$$\hat{y}_{s,k} = \hat{h}_k x_k.$$

The residual is

$$e_k = y_k - \hat{h}_k x_k.$$

If the channel estimate is accurate,

$$e_k \approx i_k + n_k.$$

The desired-signal power can be estimated as

$$\hat{P}_S = \frac{1}{N} \sum_k |\hat{h}_k x_k|^2,$$

and the interference-plus-noise power as

$$\hat{P}_{I+N} = \frac{1}{N} \sum_k |e_k|^2.$$

Then

$$\boxed{\widehat{\text{SINR}} = \frac{\hat{P}_S}{\hat{P}_{I+N}}.}$$

This is a practical and widely useful interpretation of SINR estimation.

5.14 Effective SINR

The residual

$$e_k$$

does not contain only interference and thermal noise.
It may also contain errors caused by

- imperfect channel estimation,
- phase noise,
- frequency offset,

- timing error,
- clipping,
- nonlinear distortion,
- quantization.

Thus the measured quantity is often better interpreted as an effective SINR.
This can be useful because all of these impairments affect the decoder.
A receiver does not necessarily care whether an error came from thermal noise or phase noise.
It cares whether the transmitted information can be recovered reliably.

5.15 SINR in OFDM

In an OFDM system, channel quality can vary significantly from one subcarrier to another.
The received signal on subcarrier k is

$$Y[k] = H[k]X[k] + I[k] + N[k].$$

The SINR may therefore be different for every subcarrier:

$$\gamma_k = \frac{P_S[k]}{P_I[k] + P_N[k]}.$$

Thus the receiver may observe

$$\gamma_0, \gamma_1, \ldots, \gamma_{N-1}.$$

Some subcarriers may have excellent quality while others fall into deep fades.
This becomes important when one codeword spans many subcarriers.
A simple arithmetic average does not always predict decoder performance accurately.
For this reason, many systems derive an effective SINR from the collection

$$\{\gamma_k\}.$$

This effective value is then used for link adaptation.

5.16 Relationship to EVM

Another practical quality metric is Error Vector Magnitude.
Suppose X_k is the ideal transmitted constellation point and

$$\hat{X}_k$$

is the received point after equalization.
The error vector is

$$E_k = \hat{X}_k - X_k.$$

The normalized EVM is

$$\mathrm{EVM}^2 = \frac{\mathbb{E}\left[|\hat{X}_k - X_k|^2\right]}{\mathbb{E}\left[|X_k|^2\right]}.$$

Under suitable assumptions,

$$\mathrm{SNR} \approx \frac{1}{\mathrm{EVM}^2}.$$

Therefore,

$$\mathrm{SNR}_{\mathrm{dB}} \approx -20\log_{10}(\mathrm{EVM}).$$

EVM is useful because it directly measures how far received symbols are from their ideal constellation locations.

In practice, it often reflects more impairments than thermal noise alone.

5.17 SINR and Modulation

Higher-order modulation requires better signal quality.

For example,

QPSK

contains widely separated constellation points.

By contrast,

256-QAM

packs many points into the same signal space.

As the points become closer together, smaller errors can cause the receiver to confuse neighboring symbols.

Thus,

$$\boxed{\text{higher modulation order} \rightarrow \text{greater required SINR.}}$$

A link with modest SINR may support QPSK reliably but fail badly with 256-QAM.

This is why SINR is one of the primary inputs to modulation and coding selection.

5.18 SINR and Throughput

Increasing SINR generally allows the radio to use a higher-rate modulation and coding scheme.

Conceptually,

$$\text{higher SINR} \rightarrow \text{higher MCS} \rightarrow \text{higher PHY rate.}$$

However, SINR is not the final performance metric.

The system ultimately cares about useful delivered data.

A high-rate MCS with frequent decoding failures may produce less goodput than a more conservative MCS.

Thus a practical control loop is

$$\boxed{\text{SINR} \rightarrow \text{MCS prediction} \rightarrow \text{observed BLER} \rightarrow \text{MCS adjustment.}}$$

SINR therefore predicts link capability, while decoder performance verifies that prediction.

5.19 Practical Interpretation

The three metrics can be understood through a simple analogy.

Suppose two people are trying to speak in a crowded room.

RSSI asks

> How loud is the room?

SNR asks

> How loud is the speaker compared with the background hiss?

SINR asks

> How loud is the speaker compared with everyone else talking and the background noise?

For a real communication link, the final question is usually the most relevant.

5.20 Summary

RSSI measures total received power:

$$\boxed{P_{\text{RSSI}} \approx P_S + P_I + P_N.}$$

SNR compares desired signal with noise:

$$\boxed{\text{SNR} = \frac{P_S}{P_N}.}$$

SINR compares desired signal with both interference and noise:

$$\boxed{\text{SINR} = \frac{P_S}{P_I + P_N}.}$$

The practical distinction is

$$\boxed{\text{RSSI measures strength}}$$

while

$$\boxed{\text{SINR measures usable signal quality.}}$$

RSSI is useful for rapid power measurements and coarse beam screening.

SINR is more useful for

- comparing candidate beams,

- predicting decoder performance,
- choosing modulation and coding,
- estimating achievable throughput.

The next important question is therefore how a real receiver can estimate SINR reliably from noisy observations.

That leads naturally to practical SINR estimation using pilots, residual error, EVM, and interference measurements.

Chapter 6

Practical SINR Estimation

SINR is one of the most useful measures of wireless link quality, but it is not usually available as a direct measurement.

A receiver observes a mixture of

$$\text{desired signal} + \text{interference} + \text{noise}.$$

It does not normally have separate instruments measuring each component.

The practical problem is therefore

$$\boxed{\text{estimate the desired-signal power}}$$

and

$$\boxed{\text{estimate everything that corrupts it.}}$$

The resulting estimate is then used for tasks such as

- beam selection,
- link adaptation,
- modulation and coding selection,
- interference monitoring,
- receiver diagnostics.

This chapter develops the main practical approaches to SINR estimation.

6.1 The Basic Estimation Problem

Consider the received signal

$$y = hx + i + n,$$

where

- x is the desired transmitted signal,

- h is the desired channel,
- i is interference,
- n is receiver and thermal noise.

The quantity of interest is

$$\text{SINR} = \frac{P_S}{P_I + P_N}.$$

The difficulty is that the receiver only observes

$$y.$$

It cannot directly look at the ADC output and separately observe

$$P_S, \qquad P_I, \qquad P_N.$$

The receiver must infer these quantities from additional structure in the signal. The most common sources of information are

- known pilot symbols,
- known preambles,
- unused or silent intervals,
- equalized symbol errors,
- antenna-array statistics,
- decoder feedback.

6.2 Why Known Symbols Are Useful

Suppose the receiver knows exactly what symbol should have been transmitted.

Let that known symbol be

$$x_k.$$

The received observation is

$$y_k = h_k x_k + i_k + n_k.$$

If the receiver can estimate the channel,

$$\hat{h}_k,$$

then it can predict what the desired received signal should look like:

$$\hat{y}_{s,k} = \hat{h}_k x_k.$$

This creates a powerful practical idea:

$$\boxed{\text{received signal} - \text{predicted desired signal} = \text{residual error.}}$$

The residual is

$$e_k = y_k - \hat{h}_k x_k.$$

If the channel estimate is accurate,

$$e_k \approx i_k + n_k.$$

This provides a way to estimate both parts of the SINR expression.

6.3 Pilot-Based SINR Estimation

Suppose there are N known pilot symbols.

The desired-signal power can be estimated as

$$\hat{P}_S = \frac{1}{N} \sum_{k=1}^{N} |\hat{h}_k x_k|^2.$$

The interference-plus-noise power can be estimated from the residual:

$$\hat{P}_{I+N} = \frac{1}{N} \sum_{k=1}^{N} |y_k - \hat{h}_k x_k|^2.$$

The estimated SINR is then

$$\boxed{\widehat{\text{SINR}} = \frac{\hat{P}_S}{\hat{P}_{I+N}}.}$$

In decibels,

$$\widehat{\text{SINR}}_{\text{dB}} = 10 \log_{10} \left(\widehat{\text{SINR}} \right).$$

This is one of the most practical ways to think about SINR estimation.

The receiver predicts the signal that should have arrived and treats the unexplained remainder as impairment.

6.4 A Simple Numerical Example

Suppose a pilot sequence produces an estimated desired-signal power of

$$\hat{P}_S = 10^{-5} \text{ W}.$$

Suppose the average residual power is

$$\hat{P}_{I+N} = 10^{-7} \text{ W}.$$

Then

$$\widehat{\mathrm{SINR}} = \frac{10^{-5}}{10^{-7}} = 100.$$

In decibels,

$$\widehat{\mathrm{SINR}}_{\mathrm{dB}} = 10\log_{10}(100) = 20 \text{ dB}.$$

The important point is that neither quantity needed to be measured directly at the antenna. They were inferred from

$$\text{known transmission} + \text{channel estimate} + \text{received samples}.$$

6.5 Channel Estimation Comes First

Pilot-based SINR estimation depends strongly on the quality of the channel estimate.

Suppose

$$y_k = h_k x_k + i_k + n_k.$$

The receiver subtracts

$$\hat{h}_k x_k.$$

Then

$$e_k = y_k - \hat{h}_k x_k.$$

Substituting the received signal gives

$$e_k = h_k x_k - \hat{h}_k x_k + i_k + n_k.$$

Therefore,

$$\boxed{e_k = (h_k - \hat{h}_k)x_k + i_k + n_k.}$$

This is important.

The residual contains not only interference and noise but also channel-estimation error.

If

$$\hat{h}_k$$

is inaccurate, the receiver may interpret part of the desired signal as interference.

The estimated SINR will then be lower than the true RF SINR.

6.6 Effective SINR

In practice, several effects may appear in the residual:

- thermal noise,
- external interference,
- channel-estimation error,
- timing error,
- carrier-frequency offset,
- phase noise,
- IQ imbalance,
- clipping,
- quantization,
- nonlinear distortion.

The resulting estimate is therefore often better interpreted as an *effective SINR*.
Conceptually,

$$\boxed{\mathrm{SINR}_{\mathrm{effective}} = \frac{\text{usable desired signal}}{\text{everything causing symbol error}}.}$$

For link adaptation, this can be more useful than a perfectly isolated RF measurement.
The decoder does not care whether an error came from thermal noise, phase noise, or imperfect equalization.
All of these effects reduce its ability to recover the transmitted data.

6.7 Measurement Averaging

A SINR estimate based on only one symbol is usually noisy.
Suppose

$$\gamma_k$$

is an instantaneous estimate.
A simple average over N observations is

$$\bar{\gamma} = \frac{1}{N}\sum_{k=1}^{N}\gamma_k.$$

A larger observation window generally reduces random estimation variance.
However, it also increases delay.
Thus,

$$\boxed{\text{long averaging window} \rightarrow \text{stable estimate}}$$

but

$$\boxed{\text{long averaging window} \rightarrow \text{slow response.}}$$

This tradeoff matters when the channel or interference changes rapidly.

6.8 Recursive Averaging

Instead of storing a large number of past measurements, a receiver may use exponential smoothing:

$$\hat{\gamma}_k = \alpha\gamma_k + (1-\alpha)\hat{\gamma}_{k-1}.$$

If

$$\alpha$$

is large, the estimate reacts quickly but remains noisy.
If

$$\alpha$$

is small, the estimate is smoother but responds more slowly.

This is useful in practical systems because link adaptation usually benefits from a relatively stable quality estimate rather than reacting to every instantaneous fluctuation.

6.9 Estimating Noise During Silent Intervals

Another useful method is to measure the receiver when the desired transmitter is known to be silent.

During such an interval,

$$y_{\text{off}} = i + n.$$

The corresponding measured power is approximately

$$P_{\text{off}} = P_I + P_N.$$

When the desired transmitter becomes active,

$$y_{\text{on}} = hx + i + n,$$

and

$$P_{\text{on}} \approx P_S + P_I + P_N.$$

Therefore,

$$P_S \approx P_{\text{on}} - P_{\text{off}}.$$

The resulting SINR estimate is

$$\boxed{\text{SINR} \approx \frac{P_{\text{on}} - P_{\text{off}}}{P_{\text{off}}}.}$$

This method is conceptually simple and can be quite effective.

6.10 The Limitation of Silent-Interval Estimation

The key assumption is that interference remains approximately constant between the two measurements.

That is,

$$P_{I,\text{on}} \approx P_{I,\text{off}}.$$

If interference changes rapidly, the subtraction becomes unreliable.
For example, suppose the receiver measures low interference during the silent interval.
A strong interferer then begins transmitting at the same time as the desired signal.
The receiver may incorrectly interpret part of the interference increase as desired-signal power.
Thus,

$$\boxed{\text{silent-interval methods require reasonably stationary interference.}}$$

6.11 Measuring Noise Without Interference

If the receiver can observe a frequency region known to contain neither the desired signal nor active interferers, it may estimate the thermal and receiver noise floor directly.

For samples

$$n[0], n[1], \ldots, n[N-1],$$

the noise power estimate is

$$\hat{P}_N = \frac{1}{N} \sum_{k=0}^{N-1} |n[k]|^2.$$

This can be useful for receiver calibration.
However, this gives

$$P_N,$$

not necessarily

$$P_I + P_N.$$

In interference-limited systems, thermal noise may be much smaller than interference.
Thus a good noise-floor estimate is not automatically a good SINR estimate.

6.12 EVM as a Practical Quality Estimate

Another common approach is based on Error Vector Magnitude.

Suppose the equalized received symbol is

$$\hat{X}_k,$$

and the ideal transmitted symbol is

$$X_k.$$

The error vector is

$$E_k = \hat{X}_k - X_k.$$

The normalized EVM is

$$\mathrm{EVM}^2 = \frac{\sum_k |\hat{X}_k - X_k|^2}{\sum_k |X_k|^2}.$$

If the dominant impairment behaves approximately like additive noise,

$$\mathrm{SNR} \approx \frac{1}{\mathrm{EVM}^2}.$$

Therefore,

$$\boxed{\mathrm{SNR}_{\mathrm{dB}} \approx -20 \log_{10}(\mathrm{EVM}).}$$

For example, if

$$\mathrm{EVM} = 0.1,$$

then

$$\mathrm{SNR} \approx \frac{1}{0.1^2} = 100,$$

or

$$20 \text{ dB}.$$

6.13 Why EVM Is Useful

EVM is practical because it directly measures how far the received constellation is from the ideal constellation.

Suppose the ideal QAM symbol is

$$X_k.$$

The actual received symbol might be displaced because of

- noise,
- interference,

- residual channel error,
- phase noise,
- frequency offset,
- amplifier distortion.

All of these effects increase

$$|\hat{X}_k - X_k|.$$

Thus EVM naturally captures many impairments that affect actual demodulation performance.

6.14 The Limitation of EVM-Based SINR

The approximation

$$\text{SNR} \approx \frac{1}{\text{EVM}^2}$$

is not universally exact.

It works best when the residual distortion behaves approximately like additive uncorrelated noise.

If the receiver has strong systematic errors, such as

- residual carrier offset,
- nonlinear distortion,
- imperfect equalization,
- structured interference,

then EVM may no longer map cleanly to physical SNR.

Nevertheless, it can still be a useful measure of effective link quality.

6.15 SINR Estimation in OFDM

In OFDM, the received signal is separated into subcarriers.

For subcarrier k,

$$Y[k] = H[k]X[k] + I[k] + N[k].$$

If $X[k]$ is a known pilot, the receiver estimates

$$\hat{H}[k].$$

The residual is

$$E[k] = Y[k] - \hat{H}[k]X[k].$$

The desired-signal power can be estimated as

$$\hat{P}_S[k] = |\hat{H}[k]X[k]|^2,$$

and the residual power as

$$\hat{P}_{I+N}[k] = |E[k]|^2.$$

This gives a subcarrier-level estimate

$$\boxed{\hat{\gamma}_k = \frac{|\hat{H}[k]X[k]|^2}{|E[k]|^2}.}$$

In practice, several pilots are usually averaged together because a single residual sample is too noisy to produce a stable estimate.

6.16 Frequency-Selective SINR

One of the important properties of OFDM is that SINR is not necessarily constant across frequency.

The receiver may observe

$$\gamma_0, \gamma_1, \gamma_2, \ldots, \gamma_{N-1}.$$

For example,

Subcarrier	SINR
1	28 dB
2	25 dB
3	6 dB
4	4 dB
5	27 dB

The low-SINR carriers may correspond to frequency-selective fading or narrowband interference.

This means that one scalar SINR value may not fully describe the channel.

6.17 Why Simple Averaging Can Be Misleading

Suppose half the subcarriers have

30 dB

SINR and the other half have

0 dB.

An arithmetic average in dB gives

15 dB.

However, a channel with all subcarriers near

15 dB

is not necessarily equivalent to a channel where half are excellent and half are nearly unusable. This is because FEC decoding depends on how reliability is distributed across the codeword. Thus practical systems often compute an *effective SINR*.

Conceptually,

$$\boxed{\gamma_{\text{eff}} = f(\gamma_1, \gamma_2, \ldots, \gamma_N).}$$

The function f attempts to produce a single number that predicts decoder performance more accurately than a simple mean.

6.18 SINR After Beamforming

With multiple receive antennas, SINR depends on the beamforming vector.

Suppose

$$\mathbf{y} = \mathbf{h}_s x + \mathbf{i} + \mathbf{n}.$$

The receiver applies beamformer

$$\mathbf{w}$$

to obtain

$$z = \mathbf{w}^H \mathbf{y}.$$

The desired component becomes

$$z_s = \mathbf{w}^H \mathbf{h}_s x.$$

If the transmitted desired signal has power

$$P_x,$$

then desired output power is

$$P_S = P_x \left| \mathbf{w}^H \mathbf{h}_s \right|^2.$$

Let the interference-plus-noise covariance be

$$\mathbf{R}_{I+N}.$$

Then the corresponding output power is

$$P_{I+N} = \mathbf{w}^H \mathbf{R}_{I+N} \mathbf{w}.$$

Therefore,

$$\boxed{\text{SINR}(\mathbf{w}) = \frac{P_x \left| \mathbf{w}^H \mathbf{h}_s \right|^2}{\mathbf{w}^H \mathbf{R}_{I+N} \mathbf{w}}.}$$

This equation is fundamental in multi-antenna receivers.

6.19 Why Beamforming Changes SINR

A receive beam changes both parts of the SINR.

It changes desired-signal power through

$$\left|\mathbf{w}^H \mathbf{h}_s\right|^2,$$

and it changes interference power through

$$\mathbf{w}^H \mathbf{R}_{I+N} \mathbf{w}.$$

Therefore one beam may produce

high signal power

but also

high interference power.

Another beam may produce slightly less desired power while suppressing interference much more strongly.

This is why the best beam is often the one with the best

$$\boxed{\text{post-beamforming SINR}}$$

rather than the largest RSSI.

6.20 Interference Covariance

In an antenna array, interference is not characterized only by total power.

Its spatial structure also matters.

Suppose the interference-plus-noise vector is

$$\mathbf{v} = \mathbf{i} + \mathbf{n}.$$

Its covariance matrix is

$$\boxed{\mathbf{R}_{I+N} = \mathbb{E}\left[\mathbf{v}\mathbf{v}^H\right].}$$

The diagonal elements describe power at individual antennas.

The off-diagonal elements describe correlation between antennas.

These correlations contain spatial information.

For example, a strong interferer arriving from one direction may produce a highly correlated pattern across the array.

6.21 Estimating Interference Covariance

If the desired transmitter is silent, the receiver observes

$$\mathbf{y}_k = \mathbf{i}_k + \mathbf{n}_k.$$

The covariance can be estimated using

$$\boxed{\hat{\mathbf{R}}_{I+N} = \frac{1}{N}\sum_{k=1}^{N} \mathbf{y}_k \mathbf{y}_k^H.}$$

If the desired signal is present but known, the receiver can first subtract its estimate:

$$\mathbf{e}_k = \mathbf{y}_k - \hat{\mathbf{h}}_s x_k.$$

Then

$$\boxed{\hat{\mathbf{R}}_{I+N} = \frac{1}{N}\sum_{k=1}^{N} \mathbf{e}_k \mathbf{e}_k^H.}$$

This estimate can then be used to evaluate candidate beams.

6.22 Why Covariance Is More Useful Than Scalar Power

Suppose two interferers have the same total power.

One interferer is distributed evenly across many directions.

The other arrives strongly from one specific direction.

A scalar quantity

$$P_I$$

may be identical in both cases.

However, a beamforming receiver may be able to place a spatial null toward the directional interferer.

The covariance matrix contains the information needed to identify that spatial structure.

Thus,

$$\boxed{\text{scalar interference power} = \text{how much interference exists}}$$

while

$$\boxed{\text{interference covariance} = \text{how that interference appears across the array.}}$$

6.23 Gain State and SINR Estimation

SINR estimation must also account for receiver gain.

Suppose two measurements are taken with different gain settings.

The raw digital powers cannot necessarily be compared directly.

If the analog gain changes by

$$\Delta G,$$

the digital signal power changes even if the RF input power does not.

Therefore measurements should either

- be converted back to an input-referred scale,
- or be made after proper gain normalization.

This is especially important during beam acquisition, where AGC may change gain from one beam candidate to another.

6.24 Clipping and SINR Estimation

If the receiver clips, the SINR estimate may become unreliable.

Suppose the true received signal is

$$y = s + i + n.$$

After clipping,

$$y_{\text{clip}} = \text{clip}(s + i + n).$$

The receiver can no longer assume a linear relationship between the transmitted and received signals.

The residual

$$y_{\text{clip}} - \hat{s}$$

may therefore contain large nonlinear distortion.

A SINR estimator may interpret this distortion as interference or noise.

This is technically reasonable from the decoder's perspective, but the resulting number no longer represents only the physical RF interference environment.

Therefore clipping indicators are often useful alongside SINR measurements.

6.25 Decoder-Based Quality Information

Once packets are being decoded, the receiver gains another source of information.

It can observe

- CRC success or failure,
- ACK/NACK statistics,

- block error rate,
- decoder confidence,
- LLR magnitudes.

These measurements do not directly produce RF SINR.

However, they provide information about whether the SINR estimate is actually predicting communication performance correctly.

For example, suppose the receiver reports

$$20 \text{ dB}$$

SINR but experiences very high BLER.

Possible explanations include

- biased SINR estimation,
- interference not captured by the estimator,
- stale channel estimates,
- hardware distortion,
- overly aggressive MCS selection.

Decoder feedback can therefore be used to calibrate link adaptation.

6.26 Short-Term and Long-Term SINR

It is useful to distinguish instantaneous and averaged SINR.

An instantaneous estimate describes the channel over a short interval:

$$\gamma_{\text{inst}}.$$

A long-term estimate may be

$$\gamma_{\text{avg}}.$$

These serve different purposes.

Short-term SINR may be useful for

- detecting rapid interference,
- tracking fading,
- identifying sudden beam degradation.

Longer-term SINR may be more useful for

- MCS selection,
- stable beam ranking,
- network statistics.

A practical system may maintain both.

6.27 Bias and Variance

Any SINR estimator has two important properties.

The first is bias.

An estimator is biased if its expected output is systematically different from the true value.

For example,

$$\mathbb{E}[\widehat{\text{SINR}}] \neq \text{SINR}.$$

The second is variance.

Even an unbiased estimator may fluctuate significantly from one measurement to another.

A practical receiver must therefore balance

$$\boxed{\text{bias} \quad \text{and} \quad \text{variance.}}$$

Longer averaging tends to reduce variance.

Better channel estimation tends to reduce bias.

Calibration can reduce hardware-dependent measurement error.

6.28 A Practical Estimation Hierarchy

A real receiver often does not use only one signal-quality metric.

Different metrics may be used at different stages.

During initial acquisition:

$$\boxed{\text{RSSI} + \text{pilot correlation}}$$

may be sufficient to identify candidate beams.

After synchronization:

$$\boxed{\text{pilot residual} \rightarrow \text{SINR estimate}}$$

becomes available.

After equalization:

$$\boxed{\text{EVM}}$$

provides additional information.

After decoding:

$$\boxed{\text{BLER and ACK/NACK}}$$

provide final confirmation of link quality.

A useful practical hierarchy is therefore

$$\boxed{\text{RSSI} \rightarrow \text{correlation} \rightarrow \text{SINR/EVM} \rightarrow \text{decoder statistics.}}$$

Each stage uses increasingly detailed information about the received signal.

6.29 SINR Estimation During Beam Search

Beam search introduces an additional practical constraint: estimation time.

Suppose there are

$$N_b$$

candidate beams.

If each beam requires

$$T_{\text{SINR}}$$

seconds to obtain a reliable SINR estimate, then

$$T_{\text{scan}} \approx N_b T_{\text{SINR}}.$$

A highly accurate SINR estimator may therefore be too expensive to run on every beam.

A practical solution is to use a two-stage process.

First,

$$\text{fast RSSI or correlation measurement}$$

eliminates poor candidates.

Then,

$$\text{accurate SINR measurement}$$

is performed only on the strongest candidates.

Thus,

$$\boxed{\text{cheap coarse metric} \rightarrow \text{candidate reduction} \rightarrow \text{accurate quality estimate.}}$$

6.30 SINR Estimation for Link Adaptation

Once the link is operating, SINR is commonly used to select the modulation and coding scheme.

Conceptually,

$$\widehat{\text{SINR}} \rightarrow \text{MCS selection.}$$

A higher SINR may support

$$\text{higher-order modulation}$$

and

$$\text{higher coding rate.}$$

A lower SINR may require

$$\text{lower-order modulation}$$

and

stronger coding.

However, the SINR estimate is only a prediction.
The system should also observe actual decoding performance:

$$\text{SINR} \rightarrow \text{MCS} \rightarrow \text{BLER}.$$

If measured BLER is consistently worse than expected, the SINR-to-MCS mapping may need adjustment.

6.31 What a Good SINR Estimator Should Do

A useful practical SINR estimator should have several properties.
It should be

- reasonably accurate,
- computationally inexpensive,
- stable enough for control decisions,
- responsive to real channel changes,
- robust to gain changes,
- robust to channel-estimation error.

These goals conflict.
More averaging improves stability but slows response.
More sophisticated estimators improve accuracy but consume more computation.
Thus practical SINR estimation is not simply a mathematical problem.
It is an engineering tradeoff between

$$\boxed{\text{accuracy} \leftrightarrow \text{latency} \leftrightarrow \text{complexity}.}$$

6.32 Fundamental Perspective

SINR estimation is fundamentally a problem of separating what the receiver expected to see from what it could not explain.
With known pilots,

$$\boxed{\text{predicted desired signal} = \hat{h}x}$$

and

$$\boxed{\text{unexplained residual} = y - \hat{h}x.}$$

The ratio between their powers provides a practical estimate of link quality.
For a single-antenna receiver,

$$\widehat{\mathrm{SINR}} \approx \frac{\text{predicted desired-signal power}}{\text{residual power}}.$$

For a multi-antenna receiver, the same idea extends to spatial statistics and interference covariance.

The most important practical lesson is that a measured SINR is not an absolute physical truth. It is an estimate whose meaning depends on

- how the signal power was estimated,
- how interference and noise were estimated,
- how accurate the channel estimate was,
- how much averaging was performed,
- whether the receiver remained linear.

For communication-system design, the most useful estimate is often the one that best predicts actual decoder performance.

This leads naturally into OFDM, where signal quality must be estimated separately across many frequency-domain subcarriers before being converted into a useful link-quality metric.

Chapter 7

Orthogonal Frequency-Division Multiplexing

Orthogonal Frequency-Division Multiplexing, or OFDM, is one of the most important modulation techniques used in modern broadband wireless communication.

Its central idea is simple:

replace one difficult wideband channel with many simpler narrowband channels.

Instead of sending one high-rate symbol stream across the entire channel bandwidth, OFDM divides the available bandwidth into many closely spaced subcarriers.

Each subcarrier carries a lower-rate data stream.

Because each subcarrier occupies only a small portion of the total bandwidth, the channel seen by each subcarrier is often much simpler than the original wideband channel.

This makes OFDM especially effective in environments with multipath and frequency-selective fading.

7.1 Motivation

Suppose a communication system has total available bandwidth

$$B.$$

A conventional wideband transmission may use the entire bandwidth for one high-rate signal.

The problem is that a wideband signal may experience substantially different channel gain and phase across different frequencies.

The channel is therefore described by

$$H(f),$$

rather than by a single complex coefficient.

This produces frequency-selective distortion.

OFDM approaches the problem differently.

The total bandwidth is divided into

$$N$$

subcarriers with frequencies

$$f_0, f_1, \ldots, f_{N-1}.$$

Each subcarrier has a much smaller bandwidth than the original signal.

If the subcarrier bandwidth is sufficiently narrow, then the channel over one subcarrier can often be approximated as constant:

$$H(f) \approx H[k]$$

over the bandwidth of subcarrier k.

Thus each subcarrier behaves approximately like a flat-fading narrowband channel.

This is the central motivation for OFDM.

7.2 Parallel Transmission

Suppose a transmitter must send a high-rate stream of bits.

Instead of transmitting them sequentially through one carrier, the transmitter groups them into multiple parallel streams.

Each group is mapped to a modulation symbol such as

$$\text{QPSK}, \qquad \text{16-QAM}, \qquad \text{64-QAM},$$

or another modulation format.

Let the symbols assigned to the subcarriers be

$$X[0], X[1], \ldots, X[N-1].$$

Each symbol modulates a different subcarrier.

Thus OFDM performs parallel transmission:

$$\boxed{\text{one high-rate stream} \rightarrow \text{many lower-rate subcarrier streams.}}$$

Because the individual symbol rates are lower, each subcarrier symbol lasts longer in time.

Longer symbols are less sensitive to multipath delay spread.

7.3 Subcarrier Spacing

Suppose the useful OFDM symbol duration is

$$T.$$

The subcarrier spacing is normally chosen as

$$\boxed{\Delta f = \frac{1}{T}.}$$

The subcarrier frequencies may therefore be written as

$$f_k = f_0 + k\Delta f.$$

Thus,

$$f_k = f_0 + \frac{k}{T}.$$

This particular spacing is what creates orthogonality between the subcarriers over the interval T.

7.4 Orthogonality

Consider subcarrier k:

$$s_k(t) = e^{j2\pi f_k t}.$$

Consider another subcarrier m:

$$s_m(t) = e^{j2\pi f_m t}.$$

Their inner product over one OFDM symbol is

$$\int_0^T s_k(t) s_m^*(t)\, dt.$$

Substituting,

$$\int_0^T e^{j2\pi(f_k - f_m)t}\, dt.$$

Since

$$f_k - f_m = \frac{k-m}{T},$$

the integral becomes

$$\int_0^T e^{j2\pi(k-m)t/T}\, dt.$$

For

$$k \neq m,$$

this evaluates to zero:

$$\boxed{\int_0^T e^{j2\pi f_k t} e^{-j2\pi f_m t}\, dt = 0.}$$

This is the mathematical meaning of orthogonality.

The subcarriers may overlap in frequency, but when observed over the correct symbol interval, they do not interfere with one another.

7.5 Why Overlapping Subcarriers Are Useful

A conventional frequency-division system could separate channels using large guard bands.

For example,

$$\text{carrier} \quad \text{guard band} \quad \text{carrier} \quad \text{guard band}.$$

This prevents interference but wastes spectrum.

OFDM allows the spectra of neighboring subcarriers to overlap.

The orthogonality condition ensures that the peak of one subcarrier occurs where the other subcarriers contribute zero interference at the FFT sampling points.

Thus OFDM achieves high spectral efficiency.

Conceptually,

$$\boxed{\text{spectral overlap} + \text{mathematical orthogonality} = \text{efficient frequency use.}}$$

7.6 The OFDM Transmitter

The transmitter begins with information bits.

These are mapped to complex modulation symbols:

$$X[0], X[1], \ldots, X[N-1].$$

The time-domain OFDM waveform is then generated by the inverse discrete Fourier transform.

For sample index n,

$$\boxed{x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] e^{j2\pi kn/N}.}$$

This equation shows that the time-domain waveform is the sum of all subcarriers.

Each $X[k]$ determines the amplitude and phase of one subcarrier.

In practice, this operation is implemented efficiently using the Inverse Fast Fourier Transform, or IFFT.

Thus the transmitter processing chain is approximately

$$\boxed{\text{bits} \rightarrow \text{QAM symbols} \rightarrow \text{subcarrier allocation} \rightarrow \text{IFFT} \rightarrow \text{time-domain samples.}}$$

7.7 The OFDM Receiver

The receiver performs the inverse process.

After sampling and synchronization, the receiver collects one OFDM symbol:

$$y[0], y[1], \ldots, y[N-1].$$

It applies the Fast Fourier Transform:

$$Y[k] = \sum_{n=0}^{N-1} y[n] e^{-j2\pi kn/N}.$$

This separates the composite time-domain waveform back into individual frequency-domain subcarriers.

The receiver chain is therefore

$$\boxed{\text{received samples} \rightarrow \text{FFT} \rightarrow \text{subcarrier symbols} \rightarrow \text{equalization} \rightarrow \text{demodulation} \rightarrow \text{bits.}}$$

The FFT and IFFT make OFDM computationally practical.

7.8 OFDM and Multipath

Consider a multipath channel with impulse response

$$h(t).$$

The received waveform is

$$y(t) = h(t) * x(t) + n(t).$$

In the frequency domain,

$$Y(f) = H(f)X(f) + N(f).$$

Because

$$H(f)$$

may vary significantly across a wide channel, a single wideband symbol can be strongly distorted.

OFDM divides the channel into many small frequency regions.

For subcarrier k,

$$\boxed{Y[k] = H[k]X[k] + I[k] + N[k].}$$

If each subcarrier is narrow enough,

$$H(f)$$

is approximately constant across that subcarrier.

The receiver therefore needs only to correct one complex gain per subcarrier.

7.9 Flat Fading on Each Subcarrier

Suppose the coherence bandwidth of the channel is

$$B_c.$$

If the subcarrier bandwidth

$$\Delta f$$

is much smaller than the coherence bandwidth,

$$\Delta f \ll B_c,$$

then each subcarrier experiences approximately flat fading.

Thus

$$H(f) \approx H[k]$$

within subcarrier k.

This condition is central to OFDM.

A frequency-selective wideband channel is converted into many approximately flat channels:

$$\boxed{H(f) \to H[0], H[1], \ldots, H[N-1].}$$

7.10 Frequency-Domain Equalization

Suppose the receiver knows or estimates

$$H[k].$$

The received symbol is

$$Y[k] = H[k]X[k] + N[k].$$

A simple zero-forcing equalizer computes

$$\boxed{\hat{X}[k] = \frac{Y[k]}{\hat{H}[k]}.}$$

If

$$\hat{H}[k] \approx H[k],$$

then

$$\hat{X}[k] \approx X[k] + \frac{N[k]}{H[k]}.$$

The channel amplitude and phase distortion are approximately removed.

The major advantage is that equalization is performed independently for each subcarrier.

Instead of designing a complicated long time-domain equalizer, the receiver performs one complex correction per frequency bin.

7.11 The Problem with Deep Fades

Zero-forcing equalization has a weakness.

If

$$|H[k]|$$

is very small, then

$$\frac{1}{H[k]}$$

becomes very large.

As a result, noise is strongly amplified.

Thus,

$$\hat{X}[k] = \frac{Y[k]}{H[k]}$$

can become unreliable on deeply faded subcarriers.

This motivates more robust equalizers such as MMSE equalization, which avoid fully inverting very poor channel conditions.

The important fundamental point is

equalization can remove channel distortion, but it cannot create information lost in a deep fade.

7.12 Intersymbol Interference

Multipath causes delayed copies of a symbol to arrive after the original transmission.

Suppose one OFDM symbol ends at time

$$T.$$

A delayed propagation path may still be arriving after the next symbol begins.

This produces intersymbol interference, or ISI.

Conceptually,

multipath delay → neighboring symbols overlap → ISI.

OFDM uses a cyclic prefix to reduce this problem.

7.13 The Cyclic Prefix

Suppose the useful OFDM symbol contains samples

$$x[0], x[1], \ldots, x[N-1].$$

A cyclic prefix of length L is created by copying the final L samples and placing them before the symbol.

The transmitted sequence becomes

$$x[N-L], \ldots, x[N-1], x[0], \ldots, x[N-1].$$

The copied portion is discarded at the receiver before the FFT.

At first, copying samples may appear wasteful.

However, the cyclic prefix serves two important purposes:

1. it protects the useful symbol from delayed energy from the previous symbol,
2. it converts the channel operation into circular convolution over the FFT interval.

7.14 Cyclic Prefix and Delay Spread

Suppose the important multipath delay spread is approximately

$$T_{\text{delay}}.$$

The cyclic prefix should normally satisfy

$$\boxed{T_{\text{CP}} > T_{\text{delay}}.}$$

Then delayed copies of the previous symbol fall mainly inside the cyclic prefix rather than the useful FFT interval.

After the receiver discards the cyclic prefix, the remaining samples are largely free from inter-symbol interference from the previous OFDM symbol.

7.15 Why Circular Convolution Matters

The physical wireless channel performs linear convolution:

$$y[n] = h[n] * x[n].$$

The discrete Fourier transform has a particularly simple relationship with circular convolution.

If the cyclic prefix is sufficiently long, the useful OFDM symbol behaves as if the channel performed circular convolution.

The FFT then converts this convolution into multiplication:

$$\boxed{Y[k] = H[k]X[k].}$$

This is one of the deepest practical reasons OFDM is useful.

A complicated time-domain multipath convolution becomes simple multiplication on each frequency bin.

7.16 The Cost of the Cyclic Prefix

The cyclic prefix carries no new information.

If the useful OFDM symbol duration is

$$T$$

and the cyclic-prefix duration is

$$T_{\text{CP}},$$

then the total transmitted symbol duration is

$$T_{\text{sym}} = T + T_{\text{CP}}.$$

The fraction of time carrying useful FFT data is

$$\eta_{\text{CP}} = \frac{T}{T + T_{\text{CP}}}.$$

Thus a longer cyclic prefix improves multipath robustness but reduces spectral efficiency.

This creates the tradeoff

$$\boxed{\text{long CP} \rightarrow \text{better multipath tolerance} \rightarrow \text{more overhead.}}$$

7.17 Pilot Symbols

The receiver needs an estimate of

$$H[k]$$

to equalize each subcarrier.

To obtain this estimate, the transmitter places known symbols at selected time-frequency locations.

These known symbols are called pilots or reference symbols.

Suppose the receiver knows that

$$X[k]$$

was transmitted.

It observes

$$Y[k] = H[k]X[k] + N[k].$$

A simple channel estimate is

$$\boxed{\hat{H}[k] = \frac{Y[k]}{X[k]}.}$$

Because $X[k]$ is known, the unknown quantity is the channel.

7.18 Pilot Placement

Pilots do not necessarily occupy every subcarrier.

That would consume too much capacity.

Instead, pilots are usually distributed across

$$\text{frequency}$$

and

$$\text{time.}$$

The receiver estimates the channel at pilot positions and interpolates between them.

Conceptually,

$$\boxed{\text{pilot measurements} \rightarrow \text{channel estimates} \rightarrow \text{interpolation} \rightarrow \hat{H}[k].}$$

If the channel varies rapidly in frequency, pilots must be spaced closely enough in frequency.

If the channel varies rapidly in time, pilots must be repeated frequently enough in time.

Thus pilot spacing is related to

$$\text{coherence bandwidth}$$

and

$$\text{coherence time.}$$

7.19 Pilots and Receiver Measurements

Pilots support more than channel estimation.
They can also be used for

- phase tracking,
- frequency-offset correction,
- SINR estimation,
- timing refinement.

After estimating the desired component,

$$\hat{H}[k]X[k],$$

the receiver can form the residual

$$\boxed{E[k] = Y[k] - \hat{H}[k]X[k].}$$

This residual contains information about

- interference,
- thermal noise,
- channel-estimation error,
- synchronization error,
- receiver distortion.

Thus pilots provide a reference against which receiver quality can be measured.

7.20 Frequency-Selective Channel Quality

The channel response

$$H[k]$$

may vary substantially across frequency.
For example,

Subcarrier	$\lvert H[k]\rvert$
1	strong
2	strong
3	deep fade
4	medium
5	strong

A deep fade on one subcarrier does not necessarily imply a poor channel everywhere.
Likewise, interference may be concentrated in only part of the spectrum.
Therefore SINR may vary across frequency:

$$\gamma_0, \gamma_1, \ldots, \gamma_{N-1}.$$

For subcarrier k,

$$\gamma_k = \frac{P_S[k]}{P_I[k] + P_N[k]}.$$

This per-subcarrier view is one of the major strengths of OFDM.

7.21 Subcarrier Allocation

Because subcarrier quality differs, a communication system does not necessarily need to treat every frequency identically.

For example,

$$\text{high-SINR subcarrier}$$

may support a higher modulation order.
A weaker subcarrier may require more robust modulation.
Conceptually,

$$\text{good channel} \rightarrow \text{more bits}$$

and

$$\text{poor channel} \rightarrow \text{fewer bits or no transmission.}$$

The exact adaptation strategy depends on the system design.

Even when one MCS is used across many subcarriers, the distribution of subcarrier SINRs remains important for predicting decoder performance.

7.22 OFDM Symbol Duration

The useful symbol duration is related to subcarrier spacing:

$$T = \frac{1}{\Delta f}.$$

Therefore narrow subcarrier spacing implies long OFDM symbols:

$$\Delta f \downarrow \quad \Rightarrow \quad T \uparrow .$$

Longer symbols improve resistance to a fixed multipath delay because the delay occupies a smaller fraction of the symbol.

However, longer symbols can make the system more sensitive to time variation during the symbol.

Thus OFDM parameter selection includes a fundamental tradeoff between

multipath robustness

and

time variation.

7.23 Peak-to-Average Power Ratio

The OFDM time-domain waveform is the sum of many subcarriers:

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] e^{j2\pi kn/N}.$$

At some samples, the subcarrier phases partially cancel.

At other samples, many subcarriers align constructively.

The resulting waveform may therefore have occasional large peaks.

Peak-to-Average Power Ratio is defined as

$$\boxed{\text{PAPR} = \frac{\max_n |x[n]|^2}{\mathbb{E}[|x[n]|^2]}.}$$

A high PAPR means that the maximum instantaneous signal power may be much larger than the average power.

7.24 Why PAPR Matters

Suppose a power amplifier is configured so that the average OFDM waveform is close to its maximum output level.

A large instantaneous OFDM peak may then push the amplifier into compression.

Similarly, a receiver signal with high PAPR may exceed the ADC range even though its average power appears safe.

Therefore OFDM systems require headroom.

This affects

- transmit power amplifiers,
- ADCs,
- DACs,
- analog gain settings,

- digital fixed-point scaling.

Thus the PAPR problem directly connects OFDM with the receiver dynamic-range considerations discussed earlier.

7.25 Clipping an OFDM Signal

Suppose the time-domain OFDM signal is

$$x[n].$$

If it exceeds the hardware range, the waveform may become

$$x_{\text{clip}}[n] = \text{clip}(x[n]).$$

Because clipping is nonlinear, the resulting error is not confined to one subcarrier.
The FFT of the clipped waveform generally contains distortion across many frequency bins.
Thus,

$$\boxed{\text{time-domain clipping} \rightarrow \text{frequency-domain distortion across many subcarriers.}}$$

This can degrade EVM and increase decoder error rates even on subcarriers whose original symbols were otherwise well behaved.

7.26 Timing Synchronization

The receiver must know where the OFDM symbol begins.
Suppose the receiver places its FFT window at the wrong time.
Then the collected samples may include portions of neighboring symbols.
This can cause

- intersymbol interference,
- phase errors,
- loss of orthogonality.

The receiver therefore performs timing synchronization before or during OFDM demodulation.
Known preambles or correlation sequences are often used to identify the correct symbol boundary.

7.27 Carrier-Frequency Offset

The transmitter and receiver oscillators are not perfectly identical.
Suppose their carrier-frequency difference is

$$\Delta f_{\text{CFO}}.$$

The received signal then contains an additional rotation

$$e^{j2\pi\Delta f_{\mathrm{CFO}}t}.$$

This frequency error shifts the received subcarriers relative to the FFT bins.
The carefully designed orthogonality between neighboring subcarriers is no longer exact.
The result is inter-carrier interference, or ICI.
Conceptually,

carrier-frequency offset → subcarrier misalignment → loss of orthogonality → ICI.

7.28 Phase Error

Even after coarse frequency-offset correction, the received constellation may experience phase rotation.

For a subcarrier symbol,

$$Y[k]$$

may be rotated relative to its expected constellation position.
The receiver must estimate and correct this phase.
Pilots are commonly used to track the remaining phase error.

This is especially important for higher-order modulation, where small phase errors can move a symbol closer to an incorrect constellation point.

7.29 Inter-Carrier Interference

In ideal OFDM,

$$\text{subcarrier } k$$

does not interfere with

$$\text{subcarrier } m$$

when

$$k \neq m.$$

If synchronization is imperfect, this condition no longer holds.
The received symbol may then take the more general form

$$Y[k] = H[k]X[k] + \sum_{m\neq k} C_{k,m}X[m] + N[k],$$

where

$$C_{k,m}$$

represents leakage from other subcarriers.
The second term is inter-carrier interference.
Thus maintaining orthogonality is essential to OFDM performance.

7.30 OFDM and SINR Estimation

OFDM provides a natural framework for measuring link quality in frequency.

For subcarrier k,

$$Y[k] = H[k]X[k] + I[k] + N[k].$$

If the receiver estimates

$$\hat{H}[k]$$

and knows

$$X[k],$$

then

$$E[k] = Y[k] - \hat{H}[k]X[k]$$

provides an estimate of residual impairment.

A practical per-subcarrier SINR estimate can therefore be formed from

$$\hat{P}_S[k] = |\hat{H}[k]X[k]|^2$$

and

$$\hat{P}_{I+N}[k] \approx |E[k]|^2.$$

Thus,

$$\hat{\gamma}_k = \frac{\hat{P}_S[k]}{\hat{P}_{I+N}[k]}.$$

Multiple pilot observations are usually averaged to obtain a more stable estimate.

7.31 OFDM and Beamforming

With multiple antennas, OFDM and beamforming can be combined.

For subcarrier k, the antenna-array signal may be written as

$$\mathbf{Y}[k] = \mathbf{h}[k]X[k] + \mathbf{I}[k] + \mathbf{N}[k].$$

A beamformer applies

$$\mathbf{w}[k]$$

to produce

$$Z[k] = \mathbf{w}^H[k]\mathbf{Y}[k].$$

The effective desired channel becomes

$$\mathbf{w}^H[k]\mathbf{h}[k].$$

This shows that the spatial channel can also vary across frequency.
Depending on the architecture, a system may use

- one beam across many subcarriers,
- different digital weights across frequency,
- combinations of analog and digital beamforming.

The important point is that OFDM separates the frequency dimension while beamforming processes the spatial dimension.

7.32 The Fundamental OFDM Tradeoffs

OFDM provides powerful advantages, but they come with several design tradeoffs.
A larger number of subcarriers produces narrower subcarrier bandwidth:

$$N \uparrow \rightarrow \Delta f \downarrow .$$

This makes each subcarrier more nearly flat fading.
However, it also produces longer OFDM symbols and may increase sensitivity to time variation.
A longer cyclic prefix improves multipath protection:

$$T_{\mathrm{CP}} \uparrow \rightarrow \text{greater delay tolerance.}$$

But it also increases overhead.
Greater signal headroom reduces clipping:

$$\text{headroom} \uparrow \rightarrow \text{clipping probability} \downarrow .$$

But it reduces power-amplifier and ADC utilization.
Thus OFDM design involves balancing

frequency selectivity, delay spread, time variation, overhead, dynamic range.

7.33 Practical OFDM Processing Chain

A simplified transmitter can be summarized as

bits → FEC → QAM mapping → subcarrier mapping → IFFT → cyclic prefix → DAC/RF.

The receiver performs approximately the reverse:

RF/ADC → synchronization → remove CP → FFT → channel estimation → equalization → QAM detection →

This processing chain connects OFDM directly with the other topics developed in the preceding chapters.

7.34 Fundamental Perspective

The central reason OFDM is effective is not simply that it uses many carriers.

Its real advantage is that it transforms the structure of the channel.

A broadband multipath channel begins as

$$\boxed{y(t) = h(t) * x(t) + n(t),}$$

where equalization may require handling many delayed paths.

After OFDM processing, the problem becomes approximately

$$\boxed{Y[k] = H[k]X[k] + I[k] + N[k]}$$

for each subcarrier.

Thus,

$$\boxed{\text{one difficult convolution problem} \rightarrow \text{many simple complex multiplications.}}$$

The cyclic prefix protects this structure against multipath.

The FFT separates the subcarriers.

Pilots estimate the channel.

Equalization removes the per-subcarrier channel distortion.

SINR measurements describe the quality of each subcarrier.

The remaining challenge is deciding how aggressively to transmit data over the channel.

That leads naturally to modulation, coding, and link adaptation, where the receiver uses measured channel quality to determine how many useful bits can be transmitted reliably.

Chapter 8

Modulation and Coding

A wireless link must decide how much information to place into each transmitted symbol and how much redundancy to add for error protection.

These two decisions are controlled by

- modulation,
- forward error correction coding.

Modulation determines how many information bits are represented by each transmitted symbol.

Coding determines how much redundancy is added so that corrupted bits can be recovered at the receiver.

Together, these choices determine the operating point of the physical layer.

The central tradeoff is

higher data rate $\leftrightarrow$ greater sensitivity to noise and interference.

8.1 From Bits to Symbols

Digital communication begins with a sequence of bits:

$$b_0, b_1, b_2, \ldots$$

These bits must be converted into a physical waveform that can be transmitted over the radio channel.

A modulation scheme groups several bits together and maps each group to a complex-valued symbol.

For example, two bits may form one symbol:

$$00, \quad 01, \quad 10, \quad 11.$$

In complex baseband form, a transmitted symbol can be written as

$$X = I + jQ,$$

where

- I is the in-phase component,

- Q is the quadrature component.

Each allowed value of

$$I + jQ$$

corresponds to a point in the constellation.

The receiver observes a noisy version of that point and attempts to determine which transmitted symbol was most likely sent.

8.2 Constellation Size

Suppose a modulation scheme contains

$$M$$

possible symbols.

Each symbol can represent

$$\boxed{\log_2 M}$$

bits.

For example,

$$M = 4$$

gives

$$\log_2 4 = 2$$

bits per symbol.

Similarly,

$$M = 16$$

gives

$$4$$

bits per symbol.

Thus,

$$\text{QPSK} \rightarrow 2 \text{ bits/symbol},$$

$$\text{16-QAM} \rightarrow 4 \text{ bits/symbol},$$

$$\text{64-QAM} \rightarrow 6 \text{ bits/symbol},$$

and

$$\text{256-QAM} \rightarrow 8 \text{ bits/symbol}.$$

Increasing the modulation order therefore increases the number of bits carried by each transmitted symbol.

8.3 QPSK

Quadrature Phase-Shift Keying, or QPSK, uses four constellation points.

A simplified constellation is

$$\{+1+j, +1-j, -1+j, -1-j\}.$$

Each point represents two bits.

For example,

$$00 \to +1+j,$$

$$01 \to -1+j,$$

$$11 \to -1-j,$$

$$10 \to +1-j.$$

The exact bit assignment may vary, but neighboring points are often assigned bit patterns that differ by only one bit. This is known as Gray coding.

QPSK is relatively robust because its constellation points are widely separated.

This makes it useful under low-SINR conditions.

8.4 Quadrature Amplitude Modulation

Higher-order modulation commonly uses Quadrature Amplitude Modulation, or QAM.

In QAM, both the in-phase and quadrature amplitudes vary.

A 16-QAM constellation contains

$$16$$

possible symbols.

Since

$$\log_2(16) = 4,$$

each symbol carries four bits.

A typical 16-QAM constellation uses several amplitude levels on each axis.

For example,

$$I, Q \in \{-3, -1, +1, +3\}.$$

The 16 possible combinations create a two-dimensional grid.

Similarly,

$$64\text{-QAM}$$

contains 64 possible points, and

$$256\text{-QAM}$$

contains 256.

8.5 Why Higher-Order Modulation Requires Better SINR

The available signal space is finite.

As more constellation points are inserted into the same region, the distance between neighboring points becomes smaller.

Conceptually,

$$\boxed{M \uparrow \rightarrow \text{constellation spacing} \downarrow .}$$

Suppose the transmitted symbol is

$$X.$$

The receiver observes

$$Y = X + E,$$

where E represents noise, interference, and other distortion.

If the error is small compared with the distance between constellation points, the receiver is likely to make the correct decision.

If the error becomes comparable with that spacing, the received point may cross a decision boundary.

Then the receiver chooses the wrong symbol.

Thus,

$$\boxed{\text{higher-order modulation} \rightarrow \text{more bits per symbol} \rightarrow \text{smaller decision margin} \rightarrow \text{higher required SINR.}}$$

8.6 Decision Regions

The receiver divides the complex plane into decision regions.

For each received symbol

$$Y,$$

it chooses the constellation point closest to the observation.

Mathematically,

$$\hat{X} = \arg\min_{X_m} |Y - X_m|^2.$$

This is a maximum-likelihood decision under simple additive Gaussian-noise assumptions.

If noise pushes the received point across a decision boundary, the symbol is decoded incorrectly.

The probability of this occurring grows as

$$\text{SINR} \downarrow$$

or as

$$M \uparrow .$$

This is the basic reason why modulation order must be adapted to channel quality.

8.7 Raw Spectral Efficiency

If each symbol carries

$$\log_2 M$$

bits, then a higher-order modulation scheme carries more raw information per transmitted symbol.

A useful quantity is spectral efficiency:

$$\eta = \frac{R}{B},$$

where

- R is data rate,
- B is bandwidth.

It is commonly expressed in

$$\text{bits/s/Hz}.$$

Ignoring coding and overhead, increasing modulation order increases spectral efficiency.

For example,

$$\text{QPSK} \rightarrow 2 \text{ bits/symbol},$$

while

$$\text{256-QAM} \rightarrow 8 \text{ bits/symbol}.$$

Thus 256-QAM can theoretically carry four times as many bits per symbol as QPSK.

However, this advantage exists only if the receiver can decode those symbols reliably.

8.8 Why Modulation Alone Is Not Enough

Even with a carefully chosen modulation scheme, some symbols will inevitably be corrupted.

Possible causes include

- thermal noise,
- interference,
- fading,
- channel-estimation error,
- phase error,
- hardware distortion.

If every symbol error directly caused unrecoverable data loss, wireless communication would be extremely fragile.

Modern systems therefore add controlled redundancy before transmission.

This is the role of Forward Error Correction.

8.9 Forward Error Correction

Forward Error Correction, or FEC, transforms the original information bits into a longer coded sequence.

Suppose the original message contains

$$K$$

information bits.

The encoder produces a codeword containing

$$N$$

bits, where

$$N > K.$$

The coding rate is

$$\boxed{R_c = \frac{K}{N}.}$$

For example, if

$$K = 500$$

information bits are encoded into

$$N = 1000$$

coded bits, then

$$R_c = \frac{500}{1000} = 0.5.$$

Half of the transmitted bits represent original information, while the remainder provide redundancy.

8.10 Coding Rate and Redundancy

The coding rate directly controls the amount of redundancy.

A low coding rate such as

$$R_c = 0.5$$

contains substantial redundancy.

A high coding rate such as

$$R_c = 0.9$$

contains much less.

Therefore,

$$\boxed{R_c \downarrow \rightarrow \text{more redundancy} \rightarrow \text{greater error protection.}}$$

Conversely,

$$\boxed{R_c \uparrow \rightarrow \text{less redundancy} \rightarrow \text{higher useful data rate.}}$$

Coding therefore creates another tradeoff between reliability and throughput.

8.11 How Error-Correcting Codes Help

The purpose of a code is to introduce mathematical structure into the transmitted sequence.

Suppose the original information bits are

$$\mathbf{b}.$$

The encoder generates

$$\mathbf{c} = \text{Encode}(\mathbf{b}).$$

The receiver observes a noisy version of the encoded sequence and uses the known structure of the code to infer the most likely original message.

Thus the receiver does not examine each bit independently.

It asks which valid codeword is most consistent with the received observations.

This allows the decoder to recover from some errors without requiring retransmission.

8.12 Common Error-Correcting Codes

Several coding families are widely used in communication systems.

Examples include

- convolutional codes,
- Turbo codes,
- LDPC codes,
- Polar codes.

Different standards choose different codes, but the fundamental objective is the same:

$$\boxed{\text{add structured redundancy so that corrupted data can be recovered.}}$$

Modern high-throughput systems frequently use powerful iterative codes such as LDPC codes.

8.13 Soft Information

A modern receiver usually does not immediately convert every received symbol into definite bits.

Instead, it estimates how confident it is about each bit.

Suppose the receiver is deciding whether a bit is

$$0$$

or

$$1.$$

A common confidence metric is the log-likelihood ratio:

$$\boxed{\mathrm{LLR}(b) = \log \frac{P(b = 1 \mid y)}{P(b = 0 \mid y)}.}$$

A large positive value indicates strong evidence for one bit value.

A large negative value indicates strong evidence for the other.

A value near zero indicates uncertainty.

The decoder can use this soft information to make better decisions than it could using only hard zeros and ones.

8.14 Pre-FEC and Post-FEC Errors

The receiver may make many incorrect raw bit decisions before error correction.

Define

$$\mathrm{BER}_{\mathrm{pre-FEC}}$$

as the bit error rate before decoding.

After FEC, the bit error rate may be much smaller:

$$\mathrm{BER}_{\mathrm{post-FEC}}.$$

A successful code can therefore produce

$$\mathrm{BER}_{\mathrm{pre-FEC}} > 0$$

while

$$\mathrm{BER}_{\mathrm{post-FEC}} \approx 0.$$

This distinction is important.

A noisy constellation does not automatically imply lost packets.

The FEC decoder may still successfully reconstruct the transmitted information.

8.15 The Coding Waterfall

Powerful error-correcting codes often exhibit a sharp transition in performance.
Above a certain SINR range, decoding succeeds with high probability.
Slightly below that region, the decoder may begin to fail frequently.
This behavior is often called the *waterfall*.
Conceptually,

$$\text{high SINR} \to \text{very low error rate},$$

while a relatively small reduction in SINR can lead to

$$\text{rapidly increasing BLER}.$$

This is why MCS selection can be sensitive to only a few decibels of SINR change.

8.16 Combining Modulation and Coding

Modulation and coding affect data rate in different ways.
Modulation controls

$$\text{bits per symbol}.$$

Coding controls

$$\text{fraction of transmitted bits that contain new information}.$$

Suppose the modulation order is M.
Each modulation symbol carries

$$\log_2 M$$

coded bits.
Only the fraction

$$R_c$$

represents original information.
Therefore the number of useful information bits per symbol is approximately

$$\boxed{R_c \log_2 M.}$$

This is a useful expression for understanding MCS efficiency.

8.17 Example of Modulation and Coding Rate

Consider QPSK with coding rate

$$R_c = \frac{1}{2}.$$

QPSK carries

2

coded bits per symbol.
Therefore the useful information carried per symbol is approximately

$$2 \times \frac{1}{2} = 1 \text{ information bit/symbol.}$$

Now consider 64-QAM with

$$R_c = \frac{3}{4}.$$

64-QAM carries

6

coded bits per symbol.
The useful information per symbol is approximately

$$6 \times \frac{3}{4} = 4.5 \text{ information bits/symbol.}$$

Thus the second configuration provides much greater raw throughput.
However, it also requires a substantially cleaner channel.

8.18 Modulation and Coding Scheme

A Modulation and Coding Scheme, or MCS, combines modulation order and coding rate into one operating mode.

For example, a system may define configurations such as

$$\text{QPSK}, \quad R_c = \frac{1}{2},$$

$$\text{16-QAM}, \quad R_c = \frac{2}{3},$$

$$\text{64-QAM}, \quad R_c = \frac{3}{4},$$

$$\text{256-QAM}, \quad R_c = \frac{5}{6}.$$

These are only illustrative examples; actual MCS tables are system specific.
The important point is that each MCS represents a different combination of

$$\boxed{\text{throughput} \leftrightarrow \text{robustness.}}$$

8.19 Low-SINR Operation

When SINR is low, received symbols contain substantial uncertainty.

The system therefore benefits from

$$\text{widely separated constellation points}$$

and

$$\text{strong error correction.}$$

A typical low-SINR operating mode might therefore use

$$\boxed{\text{QPSK + low coding rate.}}$$

The raw data rate is relatively low, but the transmission is more robust.

8.20 High-SINR Operation

When SINR is high, the receiver can distinguish closely spaced constellation points reliably.

The system may then use

$$\text{64-QAM}$$

or

$$\text{256-QAM}$$

with a higher coding rate.

Thus,

$$\boxed{\text{high SINR} \rightarrow \text{higher modulation order} + \text{less redundancy.}}$$

This increases useful throughput.

8.21 The Wrong MCS Can Reduce Throughput

It may seem that the highest MCS should always provide the best throughput.

This is not true.

Suppose one MCS provides a nominal rate of

$$500 \text{ Mbps}$$

but produces

$$40\%$$

block error rate.

A rough first estimate of useful throughput is

$$500(1 - 0.40) = 300 \text{ Mbps}.$$

Now suppose a more conservative MCS provides

$$420 \text{ Mbps}$$

with only

$$5\%$$

block error rate.
Then

$$420(1 - 0.05) = 399 \text{ Mbps}.$$

The lower nominal rate produces substantially more useful delivered data.
Thus,

$$\boxed{\text{maximum PHY rate} \neq \text{maximum goodput.}}$$

8.22 Block Error Rate

A useful measure of decoder performance is Block Error Rate, or BLER.
It is defined as

$$\boxed{\text{BLER} = \frac{N_{\text{failed blocks}}}{N_{\text{transmitted blocks}}}.}$$

A block is commonly considered failed when its final error-detection check, such as a CRC, does not pass.
BLER is closely connected to MCS selection.
If the MCS is too aggressive for the current channel,

$$\text{BLER} \uparrow .$$

If the MCS is overly conservative,

$$\text{BLER} \downarrow,$$

but available channel capacity may be wasted.

8.23 SINR-to-MCS Mapping

A practical wireless system often contains an approximate mapping

$$\text{SINR} \to \text{MCS}.$$

For example, conceptually,

$$\text{low SINR} \to \text{QPSK} + \text{strong coding},$$

$$\text{moderate SINR} \to \text{16-QAM},$$

$$\text{higher SINR} \rightarrow \text{64-QAM},$$

$$\text{very high SINR} \rightarrow \text{256-QAM}.$$

The exact transition points depend on

- modulation,
- code family,
- block size,
- receiver implementation,
- target BLER,
- hardware impairments.

Therefore there is no universal SINR threshold for a given MCS.

8.24 Why MCS Thresholds Must Be Measured

It is possible to derive theoretical symbol-error or bit-error probabilities for ideal channels.

A real receiver, however, also experiences

- imperfect channel estimation,
- phase noise,
- interference,
- quantization,
- clipping,
- synchronization error,
- implementation losses.

Therefore practical MCS thresholds are often derived from measured or simulated BLER curves.

For each MCS, one can obtain a relationship of the form

$$\text{BLER} = f(\text{SINR}).$$

A typical curve falls sharply as SINR increases.

The scheduler or link-adaptation algorithm then chooses an MCS based on the SINR required to achieve the desired BLER.

8.25 A Practical MCS Table

Conceptually, a system may contain a table like

MCS	Modulation	Coding Strength
0	QPSK	strong
1	QPSK	moderate
2	16-QAM	strong
3	16-QAM	moderate
4	64-QAM	moderate
5	64-QAM	weak
6	256-QAM	weak

As the MCS index increases, the system generally becomes more spectrally efficient but less tolerant of poor channel conditions.

The exact organization is implementation dependent.

8.26 Effective Spectral Efficiency

Ignoring protocol overhead, a useful approximation for coded spectral efficiency is

$$\boxed{\eta \approx R_c \log_2 M \quad \text{bits/s/Hz.}}$$

For example, consider 64-QAM with coding rate

$$R_c = \frac{3}{4}.$$

Then

$$\eta = \frac{3}{4} \times 6 = 4.5 \text{ bits/s/Hz.}$$

By comparison, QPSK with coding rate

$$R_c = \frac{1}{2}$$

provides

$$\eta = \frac{1}{2} \times 2 = 1 \text{ bit/s/Hz.}$$

The 64-QAM configuration therefore carries much more information through the same amount of spectrum, provided the channel is good enough to support it.

8.27 Coding Gain

Error-correcting codes allow reliable communication at lower SINR than would otherwise be possible.

This improvement is often described in terms of coding gain.

Conceptually,

$$\boxed{\text{better code} \rightarrow \text{same error rate at lower required SINR.}}$$

However, coding gain is obtained by transmitting additional redundancy and performing more decoding work.

Therefore coding has costs in

- raw throughput,
- computational complexity,
- latency.

The objective is not to maximize coding strength in isolation, but to choose enough redundancy for the current channel condition.

8.28 Interleaving

Errors in wireless channels are often clustered.

For example, a short fade may corrupt several neighboring transmitted symbols.

If adjacent coded bits all belong to the same portion of a codeword, this burst may be difficult for the decoder to correct.

Interleaving rearranges coded bits before transmission.

Neighboring coded bits can be spread across different

- OFDM subcarriers,
- OFDM symbols,
- time intervals.

At the receiver, the original ordering is restored before decoding.

This transforms a concentrated burst of channel errors into a more distributed error pattern.

Thus,

$$\boxed{\text{interleaving} \rightarrow \text{spread burst errors} \rightarrow \text{improve FEC effectiveness.}}$$

8.29 Modulation and Coding in OFDM

In an OFDM system, coded bits are mapped onto QAM symbols and then placed onto subcarriers.

The transmitter processing chain is approximately

$$\boxed{\text{information bits} \rightarrow \text{FEC encoder} \rightarrow \text{interleaver} \rightarrow \text{QAM mapper} \rightarrow \text{OFDM subcarriers.}}$$

At the receiver,

$$\boxed{\text{equalized subcarriers} \rightarrow \text{soft QAM demapper} \rightarrow \text{deinterleaver} \rightarrow \text{FEC decoder.}}$$

The demapper typically generates soft information rather than hard bits.

This soft information is passed to the FEC decoder.

8.30 Why Frequency-Selective SINR Matters

In OFDM,

$$\gamma_k$$

may vary greatly across subcarriers.

Some coded bits may therefore be transmitted through excellent channel conditions while others pass through deep fades.

Because the FEC code spans many transmitted bits, decoder performance depends on the collection

$$\gamma_0, \gamma_1, \ldots, \gamma_{N-1},$$

not necessarily on a simple arithmetic average.

This is why practical link-adaptation systems often use an effective SINR metric that attempts to predict how the FEC decoder will behave.

8.31 Modulation, Coding, and Shannon Capacity

There is a fundamental theoretical relationship between channel quality and achievable data rate.

For an ideal channel of bandwidth B and signal-to-noise ratio SNR, Shannon's capacity formula is

$$\boxed{C = B \log_2(1 + \mathrm{SNR}).}$$

This equation does not specify a particular modulation or code.

Instead, it provides an upper bound on reliable communication rate under ideal assumptions.

Practical modulation and coding schemes attempt to operate as efficiently as possible below this limit.

Higher SINR allows greater information density.

Lower SINR requires more conservative transmission.

Thus the practical MCS adaptation process reflects the same fundamental principle described by Shannon capacity.

8.32 Why Practical Systems Use Discrete MCS Levels

In theory, transmission rate could vary continuously with channel quality.

Real systems typically support a finite set of configurations.

For example,

$$\mathrm{MCS}_0, \mathrm{MCS}_1, \ldots, \mathrm{MCS}_{N-1}.$$

Each corresponds to a tested combination of modulation and coding parameters.

This simplifies

- transmitter implementation,
- receiver implementation,

- scheduler behavior,
- interoperability,
- performance validation.

The adaptation problem therefore becomes

$$\boxed{\text{choose the highest usable MCS from a discrete table.}}$$

8.33 Fundamental Tradeoff

Modulation and coding solve opposite sides of the same problem.

Higher modulation order increases the amount of information placed into each radio symbol:

$$M \uparrow \rightarrow \text{throughput} \uparrow .$$

Stronger coding adds redundancy:

$$R_c \downarrow \rightarrow \text{robustness} \uparrow .$$

The practical radio continually balances these quantities.

At poor channel quality,

$$\boxed{\text{few bits per symbol} + \text{large redundancy}}$$

is appropriate.

At good channel quality,

$$\boxed{\text{many bits per symbol} + \text{small redundancy}}$$

is appropriate.

The objective is not simply to transmit as many bits as possible.

It is to transmit as many bits as possible while keeping the probability of decoding failure sufficiently low.

8.34 Practical Perspective

The complete relationship can be summarized as

$$\boxed{\text{SINR} \rightarrow \text{MCS} \rightarrow \text{PHY rate} \rightarrow \text{BLER} \rightarrow \text{goodput.}}$$

SINR indicates what the channel may support.

The MCS determines how aggressively information is packed into the waveform.

FEC attempts to repair errors.

BLER reveals whether the chosen operating point is actually reliable.

The next problem is therefore not simply choosing a modulation order or coding rate.

It is choosing the MCS that maximizes useful delivered throughput.

That leads naturally to link adaptation and goodput optimization.

Chapter 9

Link Adaptation and Throughput Optimization

Wireless channel quality changes over time and across frequency. A fixed modulation and coding scheme is therefore rarely optimal under all conditions.

Link adaptation is the process of selecting transmission parameters, especially the modulation and coding scheme (MCS), according to current channel quality.

The goal is not simply to maximize the nominal physical-layer rate.

The goal is to maximize the amount of useful data successfully delivered.

9.1 Raw PHY Rate Versus Goodput

The physical-layer rate,

$$R_{\mathrm{PHY}},$$

describes the nominal rate at which coded information is transmitted.

However, some transmitted blocks may fail decoding.

A useful first approximation for delivered throughput is

$$\boxed{R_{\mathrm{goodput}} \approx R_{\mathrm{PHY}} \left(1 - \mathrm{BLER}\right)}$$

where

$$\mathrm{BLER} = \frac{N_{\text{failed blocks}}}{N_{\text{transmitted blocks}}}.$$

This approximation ignores several real-world effects, including

- retransmission overhead,
- protocol headers,
- scheduling overhead,
- control signaling,
- latency constraints.

Nevertheless, it captures the main principle:

$$\boxed{\text{highest nominal rate} \neq \text{highest useful throughput.}}$$

9.2 A Simple Example

Consider several possible transmission modes:

MCS	R_{PHY}	BLER	R_{goodput}
QPSK	100	0.001	99.9
16-QAM	200	0.02	196
64-QAM	350	0.15	297.5
256-QAM	500	0.60	200

Although 256-QAM provides the largest nominal rate, its high error rate causes useful throughput to fall substantially.

The 64-QAM mode provides lower raw rate but higher goodput.

This illustrates why link adaptation must consider both

$$\text{rate}$$

and

$$\text{reliability.}$$

9.3 The Link-Adaptation Objective

A simplified link-adaptation objective is

$$\boxed{\text{MCS}^* = \arg\max_{\text{MCS}} R_{\text{PHY}}(\text{MCS})\,[1 - \text{BLER}(\text{MCS}, \gamma)]}$$

where

$$\gamma$$

represents measured channel quality, often in the form of SINR.

The selected MCS should therefore be aggressive enough to provide high data rate, but not so aggressive that decoding failures dominate.

9.4 SINR as a Predictor

SINR is commonly used to predict which MCS the channel can support.

Conceptually,

$$\boxed{\text{low SINR} \rightarrow \text{robust MCS}}$$

and

$$\boxed{\text{high SINR} \rightarrow \text{higher-rate MCS.}}$$

A low-SINR channel may require

$$\text{QPSK + strong coding,}$$

while a high-SINR channel may support

$$\text{64-QAM}$$

or

$$\text{256-QAM}$$

with a higher coding rate.

The exact thresholds depend on the receiver, coding scheme, block size, and target error rate.

9.5 Effective SINR in OFDM

In OFDM, channel quality is not necessarily constant across the spectrum.

Each subcarrier may have its own SINR:

$$\gamma_0, \gamma_1, \ldots, \gamma_{N-1}.$$

A simple average,

$$\bar{\gamma} = \frac{1}{N} \sum_{k=0}^{N-1} \gamma_k,$$

may not accurately predict decoder performance.

For example, a channel in which all subcarriers have moderate SINR may behave very differently from one in which half have excellent SINR and half are in deep fades.

Practical systems therefore often compute an *effective SINR*:

$$\boxed{\gamma_{\text{eff}} = f(\gamma_0, \gamma_1, \ldots, \gamma_{N-1}).}$$

The purpose of this mapping is to summarize the frequency-selective channel with one value that better predicts whether the coded block will decode successfully.

9.6 Inner-Loop and Outer-Loop Adaptation

A practical link-adaptation system often contains two forms of feedback.

The first uses measured channel quality:

$$\gamma_{\text{eff}} \rightarrow \text{MCS selection.}$$

This provides the initial prediction.

The second uses actual decoder performance:

$$\text{ACK/NACK} \rightarrow \text{BLER estimate} \rightarrow \text{MCS correction.}$$

This second process is commonly called *outer-loop link adaptation*.

If the observed BLER is consistently too high, the system becomes more conservative.

If the observed BLER is extremely low, the system may increase the transmission rate.
Thus the complete loop is approximately

$$\boxed{\text{SINR} \rightarrow \text{MCS} \rightarrow \text{transmission} \rightarrow \text{decode} \rightarrow \text{ACK/NACK} \rightarrow \text{MCS adjustment.}}$$

9.7 Why SINR Alone Is Not Enough

SINR is an estimate of channel quality.
Actual decoding performance also depends on

- channel-estimation accuracy,
- interference structure,
- coding implementation,
- phase noise,
- clipping,
- synchronization error,
- hardware impairments.

Therefore a SINR-to-MCS mapping is not perfect.
Observed BLER provides feedback about whether the current mapping is working correctly.
In this sense,

$$\boxed{\text{SINR predicts performance}}$$

while

$$\boxed{\text{BLER measures actual performance.}}$$

9.8 Why Zero BLER Is Not Always Optimal

It may appear desirable to choose an MCS that produces essentially zero errors.
However, this can be unnecessarily conservative.
Suppose one MCS provides

$$300 \text{ Mbps}$$

with nearly zero BLER.
A slightly more aggressive MCS might provide

$$450 \text{ Mbps}$$

with a small nonzero error rate.
Even after accounting for occasional retransmissions, the second operating point may deliver more useful data.
Therefore practical systems often accept some nonzero block error rate.
The ideal operating point depends on

- retransmission cost,
- latency requirements,
- application type,
- scheduler behavior,
- channel dynamics.

9.9 Rate Adaptation Versus Reliability

The fundamental tradeoff is

$$\boxed{\text{aggressive MCS} \rightarrow \text{higher rate} + \text{higher error probability}}$$

versus

$$\boxed{\text{conservative MCS} \rightarrow \text{lower rate} + \text{greater reliability.}}$$

The purpose of link adaptation is to operate near the point where the additional rate gained from a more aggressive MCS is still worth the corresponding increase in decoding failures.

9.10 Practical Control Loop

A useful summary of the complete process is

$$\boxed{\text{measure SINR} \rightarrow \text{estimate supported MCS} \rightarrow \text{transmit} \rightarrow \text{measure BLER} \rightarrow \text{adjust MCS.}}$$

This process runs continuously.
If the channel improves,

$$\text{SINR} \uparrow,$$

the system can increase spectral efficiency.
If the channel degrades,

$$\text{SINR} \downarrow,$$

the system moves toward more robust modulation and stronger coding.

9.11 Fundamental Perspective

Link adaptation connects physical channel quality to useful application throughput.
The main quantities form the chain

$$\boxed{\text{SINR} \rightarrow \text{MCS} \rightarrow R_{\text{PHY}} \rightarrow \text{BLER} \rightarrow R_{\text{goodput}}.}$$

SINR describes the channel.
MCS determines how aggressively the channel is used.

BLER measures whether that decision was successful.
Goodput measures the final useful result.
The central objective is therefore

$$\boxed{\text{maximize successfully delivered information per unit time.}}$$

This is more important than maximizing raw PHY rate or minimizing BLER independently.

Chapter 10

Error Detection and Forward Error Correction

Wireless channels inevitably introduce errors.

Noise, interference, fading, synchronization error, and imperfect channel estimation can all cause a received symbol to differ from the transmitted symbol.

A communication system therefore needs mechanisms to

- detect errors,
- correct some errors without retransmission,
- determine when a received block is still unusable.

Forward Error Correction, or FEC, provides the first line of protection.

The transmitter intentionally adds redundancy so that the receiver can recover the original information even when some transmitted bits are corrupted.

10.1 From Information Bits to Codewords

Let the original information bits be represented by

$$\mathbf{b}.$$

An FEC encoder transforms them into a longer coded sequence

$$\boxed{\mathbf{c} = \text{Encode}(\mathbf{b}).}$$

If the information block contains

$$K$$

bits and the encoded block contains

$$N$$

bits, with

$$N > K,$$

then the coding rate is

$$R_c = \frac{K}{N}.$$

The additional

$$N - K$$

bits provide redundancy.

This redundancy introduces mathematical relationships among the transmitted bits. The decoder uses these relationships to determine which original information sequence was most likely transmitted.

10.2 Why Redundancy Helps

Suppose a codeword is transmitted through a noisy channel.

The receiver does not necessarily observe the exact transmitted sequence. Instead, some bits or symbols may be corrupted.

Without coding, these errors may directly alter the reconstructed data.

With coding, the received sequence can be compared against the set of valid codewords.

The decoder searches for a codeword that is most consistent with the received observations.

Thus,

$$\boxed{\text{redundancy} \rightarrow \text{ability to identify and correct errors.}}$$

Stronger coding generally allows the receiver to tolerate more severe channel impairment, but it reduces the fraction of transmitted bits that contain new information.

10.3 Common Forward Error Correction Codes

Several families of error-correcting codes are widely used in communication systems.

Examples include

- convolutional codes,
- Turbo codes,
- LDPC codes,
- Polar codes.

The internal mathematics of these codes differs, but their purpose is the same:

$$\boxed{\text{use structured redundancy to recover corrupted information.}}$$

Modern broadband systems often use powerful iterative codes because they can operate relatively close to the theoretical limits of reliable communication.

10.4 Symbol Errors and Bit Errors

The receiver first observes modulation symbols.

Suppose the transmitted symbol is

$$X_k$$

and the receiver estimates

$$\hat{X}_k.$$

Noise or interference may cause

$$\hat{X}_k \neq X_k.$$

This is a symbol error.

Because one modulation symbol may represent several bits, a symbol error can cause one or more bit errors.

However,

$$\boxed{\text{symbol error} \not\Rightarrow \text{packet failure.}}$$

The FEC decoder may still recover the correct information bits from the complete coded block.

10.5 Hard Decisions

A simple receiver could convert every received observation immediately into a definite bit value.

For example,

$$\hat{b} = \begin{cases} 0, & \text{if the receiver favors 0,} \\ 1, & \text{if the receiver favors 1.} \end{cases}$$

This is called a *hard decision*.

The problem is that a hard decision discards information about confidence.

A receiver may be nearly certain that a bit is

$$1,$$

or only slightly more confident in

$$1$$

than in

$$0.$$

A hard-decision representation treats both cases identically.

Modern receivers therefore usually preserve reliability information.

10.6 Soft Decisions

A soft-decision receiver estimates not only the most likely bit value, but also how confident that decision is.

A common metric is the log-likelihood ratio:

$$\boxed{\mathrm{LLR}(b) = \log \frac{P(b = 1 \mid y)}{P(b = 0 \mid y)}.}$$

The sign indicates which bit value is more likely.

The magnitude indicates confidence.

For example,

$$|\mathrm{LLR}(b)| \gg 0$$

means the receiver is relatively confident.

By contrast,

$$\mathrm{LLR}(b) \approx 0$$

indicates substantial uncertainty.

The FEC decoder uses these soft values to make a better estimate of the entire codeword.

10.7 Why Soft Decoding Helps

Suppose a codeword contains several uncertain bits.

If the decoder receives only hard decisions, it sees something like

$$1, \quad 0, \quad 1, \quad 1, \quad 0.$$

It cannot tell which bits were reliable and which were close to a decision boundary.

With soft information, the decoder might instead see conceptually

$$\text{strong } 1, \quad \text{weak } 0, \quad \text{strong } 1, \quad \text{uncertain } 1, \quad \text{strong } 0.$$

The decoder can then rely more heavily on high-confidence observations and use the code constraints to resolve uncertain bits.

This can substantially improve decoding performance.

10.8 Pre-FEC and Post-FEC Error Rates

The received coded bits may contain errors before decoding.

This is described by the pre-FEC bit error rate:

$$\mathrm{BER}_{\mathrm{pre-FEC}}.$$

After FEC decoding, many of those errors may be corrected.

The resulting error rate is

$$\mathrm{BER}_{\mathrm{post-FEC}}.$$

A successful communication system may therefore operate with

$$\mathrm{BER}_{\mathrm{pre-FEC}} > 0$$

while maintaining

$$\mathrm{BER}_{\mathrm{post-FEC}} \approx 0.$$

This is a fundamental concept in modern digital communication.

The receiver does not require every raw symbol decision to be correct.

It requires the coded block to contain enough reliable information for the decoder to reconstruct the original data.

10.9 Block-Level Decoding

FEC generally operates on groups of bits or symbols rather than isolated observations.

Suppose a coded block contains

$$N$$

coded bits.

The decoder processes the complete block and attempts to recover the original

$$K$$

information bits.

The outcome is therefore naturally evaluated at the block level.

If a block cannot be decoded correctly, it may later be counted as a block error and potentially retransmitted.

This leads to the Block Error Rate:

$$\mathrm{BLER} = \frac{N_{\text{failed blocks}}}{N_{\text{total blocks}}}.$$

BLER is often a more useful practical measure than raw BER because communication protocols typically retransmit blocks or packets rather than individual bits.

10.10 The Waterfall Region

Powerful FEC codes often exhibit a sharp change in performance as SINR varies.

At sufficiently high SINR,

$$\text{decoding succeeds almost all of the time.}$$

As SINR falls, the decoder eventually reaches a region where performance deteriorates quickly.

This region is commonly called the *waterfall region*.

Conceptually,

$$\boxed{\text{small SINR change} \rightarrow \text{large BLER change.}}$$

For example, a particular MCS may operate with

$$\text{BLER} = 1\%$$

at one SINR, but only a few decibels lower the BLER may rise to

$$20\%$$

or more.

This is why accurate link adaptation is important.

10.11 Coding Strength and Operating Point

A lower coding rate provides more redundancy.

For example,

$$R_c = 0.5$$

generally provides stronger error protection than

$$R_c = 0.9.$$

Thus,

$$\boxed{R_c \downarrow \rightarrow \text{more redundancy} \rightarrow \text{greater decoding robustness.}}$$

However, stronger coding also reduces useful information rate.

The system therefore chooses coding strength according to channel quality.

Poor channel conditions favor stronger coding.

Good channel conditions allow higher coding rates and greater throughput.

10.12 FEC and Modulation Work Together

Modulation and coding should not be considered independently.

Higher-order modulation places constellation points closer together and therefore increases the raw probability of symbol errors.

Stronger coding can compensate for some of this increased uncertainty.

Thus an MCS combines

$$\text{modulation order}$$

with

$$\text{coding rate.}$$

For example,

$$\text{QPSK + strong coding}$$

may be suitable for a weak channel, while

$$\text{256-QAM + weak coding}$$

may be suitable for a strong channel.

The physical layer therefore balances

$$\boxed{\text{information density} \leftrightarrow \text{error protection.}}$$

10.13 What Happens When FEC Is Not Enough

FEC can correct many errors, but it cannot guarantee successful decoding under arbitrarily poor conditions.

If the received block contains too little reliable information, the decoder may fail.

At that point the receiver needs a mechanism to determine whether the reconstructed block is correct.

This is typically accomplished using an error-detection code such as a Cyclic Redundancy Check, or CRC.

The basic sequence is therefore

$$\boxed{\text{received symbols} \rightarrow \text{soft demodulation} \rightarrow \text{FEC decoding} \rightarrow \text{error detection.}}$$

If the block passes the error check, it is accepted.

If the block fails, higher-layer mechanisms may request retransmission.

10.14 Fundamental Perspective

Forward error correction changes the way a receiver handles errors.

Without FEC,

$$\text{wrong symbol} \rightarrow \text{wrong data.}$$

With FEC,

$$\text{wrong symbols} \rightarrow \text{soft reliability information} \rightarrow \text{codeword inference} \rightarrow \text{possibly correct data.}$$

The central idea is

$$\boxed{\text{the receiver does not need every transmitted symbol to be correct;}}$$

$$\boxed{\text{it only needs enough reliable information to recover the original codeword.}}$$

This is why wireless systems can operate effectively even when the raw channel introduces a significant number of symbol and bit errors.

The next stage is error detection: determining whether FEC successfully reconstructed the block and, when it did not, deciding whether retransmission is required.

Chapter 11

CRC, ARQ, and HARQ

Forward error correction can recover many corrupted bits, but it cannot guarantee that every transmitted block will be decoded correctly.

A practical receiver therefore needs two additional mechanisms:

- a way to detect whether decoding succeeded,
- a way to recover data when decoding failed.

CRC provides error detection.
ARQ provides retransmission.
HARQ improves retransmission by combining information from multiple transmission attempts.

11.1 CRC Error Detection

A Cyclic Redundancy Check, or CRC, is an error-detection mechanism appended to a block before transmission.

Suppose the original information block is

$$\mathbf{b}.$$

A CRC function produces a short check value

$$r = \mathrm{CRC}(\mathbf{b}).$$

The transmitter sends both

$$\mathbf{b}$$

and

$$r.$$

After reception and FEC decoding, the receiver obtains an estimated information block

$$\hat{\mathbf{b}}$$

and an estimated CRC value

$$\hat{r}.$$

The receiver computes

$$\mathrm{CRC}(\hat{\mathbf{b}})$$

and compares it with

$$\hat{r}.$$

If they agree, the block is treated as successfully decoded.
If they do not agree, the block is considered corrupted.
Conceptually,

$$\boxed{\text{FEC attempts correction} \rightarrow \text{CRC checks the result.}}$$

11.2 Why CRC Is Still Needed After FEC

An FEC decoder may output a candidate codeword even when the received data is too corrupted to recover reliably.

The receiver therefore needs an independent check on whether the final reconstructed block is valid.

A CRC provides that check.

Thus,

$$\boxed{\text{FEC answers: “Can I reconstruct the data?”}}$$

while

$$\boxed{\text{CRC answers: “Does the reconstructed block appear valid?”}}$$

A CRC is not primarily designed to correct errors.
Its purpose is to detect them with very high probability.

11.3 Acknowledgments

Once the receiver knows whether a block was successfully decoded, it can return feedback to the transmitter.

Two common responses are

$$\mathrm{ACK}$$

and

$$\mathrm{NACK}.$$

ACK means acknowledgment:

$$\boxed{\text{ACK} \rightarrow \text{block received successfully.}}$$

NACK means negative acknowledgment:

$$\boxed{\text{NACK} \rightarrow \text{block must be recovered.}}$$

The feedback closes the loop between transmitter and receiver.

11.4 Automatic Repeat Request

Automatic Repeat reQuest, or ARQ, is the simplest retransmission mechanism.

The basic sequence is

$$\boxed{\text{transmit} \rightarrow \text{decode} \rightarrow \text{CRC check} \rightarrow \begin{cases} \text{ACK}, & \text{success}, \\ \text{NACK}, & \text{failure}. \end{cases}}$$

If the transmitter receives an ACK, it proceeds to new data.

If it receives a NACK, it retransmits the failed block.

Conceptually,

$$\text{failed packet} \rightarrow \text{send it again.}$$

ARQ is simple and reliable, but it can be inefficient because the receiver may discard useful information from the failed first transmission.

11.5 The Cost of Retransmission

Retransmissions consume resources.

Each repeated block uses

- additional airtime,
- additional spectrum,
- additional energy,
- additional latency.

If the error rate becomes high, a large portion of the available channel capacity may be spent repeating previously transmitted data.

This is one reason link adaptation attempts to avoid operating at excessively high BLER.

11.6 Hybrid ARQ

Hybrid Automatic Repeat reQuest, or HARQ, improves on basic ARQ.

The key idea is simple:

$$\boxed{\text{do not throw away information from a failed transmission.}}$$

Instead, the receiver stores soft information from the first attempt and combines it with information from later retransmissions.

Suppose the first received observation is

$$y_1 = hx + n_1.$$

If decoding fails, the receiver preserves information derived from

$$y_1.$$

A second transmission produces

$$y_2 = hx + n_2.$$

The receiver then uses both observations.
In a simplified example,

$$y_{\text{combined}} = y_1 + y_2.$$

Since the noise realizations

$$n_1$$

and

$$n_2$$

are different, combining the two observations can improve the effective reliability of the transmitted data.

11.7 Why Combining Helps

Suppose the same symbol is transmitted twice:

$$y_1 = hx + n_1,$$

$$y_2 = hx + n_2.$$

If the receiver averages them,

$$\bar{y} = \frac{y_1 + y_2}{2},$$

then

$$\bar{y} = hx + \frac{n_1 + n_2}{2}.$$

The desired component remains

$$hx,$$

while the averaged noise becomes less variable when the noise samples are independent.
Thus repeated observations can improve effective SNR.
The practical receiver usually performs this combination using soft bit information rather than by simply adding raw waveform samples, but the principle is the same.

11.8 Chase Combining

One form of HARQ is Chase combining.

The transmitter retransmits essentially the same coded symbols as before.

Suppose the original transmission contains coded sequence

$$[c_1, c_2, c_3, c_4].$$

If decoding fails, the retransmission again sends

$$[c_1, c_2, c_3, c_4].$$

The receiver now has two observations of each coded bit.

It combines their reliability information before attempting FEC decoding again.

Thus,

Chase combining = repeat the same coded information and combine observations.

The failed first transmission is therefore still useful.

11.9 Incremental Redundancy

A more flexible HARQ method is incremental redundancy.

Instead of repeating exactly the same coded bits, the transmitter sends additional parity information.

Suppose the FEC encoder conceptually produces

$$[c_1, c_2, c_3, c_4, c_5, c_6].$$

The first transmission may contain

$$[c_1, c_2, c_3, c_4].$$

If decoding fails, the next transmission may provide

$$[c_5, c_6].$$

The receiver can then decode using the larger set

$$[c_1, c_2, c_3, c_4, c_5, c_6].$$

The effective code has become stronger because more redundancy is now available.

Thus,

incremental redundancy = send new parity information after failure.

11.10 HARQ as Adaptive Coding

Incremental redundancy can be interpreted as a form of adaptive coding.

The first transmission may use a relatively high effective coding rate.

If the channel is good, decoding succeeds immediately and no additional redundancy is needed.

If the channel is poor, additional parity bits are transmitted.

The effective code rate therefore decreases after each retransmission.

Conceptually,

$$\boxed{\text{good channel} \rightarrow \text{little redundancy}}$$

while

$$\boxed{\text{poor channel} \rightarrow \text{additional redundancy.}}$$

This makes HARQ efficient because strong error protection is added only when required.

11.11 Chase Combining Versus Incremental Redundancy

The two main HARQ approaches can be summarized as follows.

Method	Retransmission content
Chase combining	repeat the same coded bits
Incremental redundancy	send additional parity bits

Chase combining improves reliability by obtaining another observation of the same information.

Incremental redundancy improves reliability by increasing the total amount of coding information available to the decoder.

Both reuse information from previous failed transmissions.

11.12 Soft Combining

HARQ normally operates on soft information.

Suppose the first transmission produces a bit log-likelihood ratio

$$L_1.$$

The retransmission produces another value

$$L_2.$$

For repeated observations under suitable assumptions, the combined information may be approximated as

$$\boxed{L_{\text{combined}} = L_1 + L_2.}$$

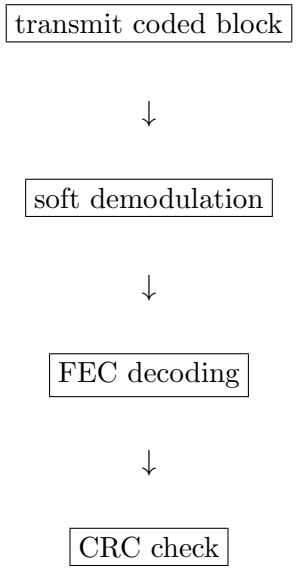
If both observations support the same bit value, the magnitude of the combined LLR increases.

The FEC decoder then operates on the improved soft information.

This is more useful than simply taking a majority vote between hard bit decisions.

11.13 HARQ Process

A practical HARQ process can be summarized as



If the CRC passes,

is returned.
If the CRC fails,

is returned and the receiver preserves useful soft information.

A later transmission is then combined with the stored information before decoding is attempted again.

11.14 Maximum Retransmissions

A system cannot retransmit the same block indefinitely.

A maximum number of HARQ attempts is usually imposed.

Suppose

$$N_{\mathrm{HARQ,max}}$$

is the maximum number of permitted transmissions.

After this limit, the block may be

- discarded,
- reported as failed to a higher layer,
- recovered by another retransmission mechanism.

This prevents a single badly corrupted block from consuming unlimited radio resources.

11.15 HARQ and Latency

HARQ improves reliability, but every additional transmission introduces delay.

Suppose one HARQ round requires

$$T_{\text{round}}.$$

After r retransmissions, the added latency may be approximately

$$T_{\text{HARQ}} \approx rT_{\text{round}},$$

ignoring other scheduling effects.

Thus HARQ creates a tradeoff between

$$\boxed{\text{reliability} \leftrightarrow \text{latency.}}$$

Applications that are highly latency sensitive may therefore prefer a more robust initial MCS rather than relying heavily on retransmissions.

11.16 HARQ and Goodput

HARQ also affects useful throughput.

If every block succeeds on the first transmission, nearly all airtime can be used for new information.

If many blocks require repeated attempts, a growing portion of the channel is spent retransmitting old information.

Therefore,

$$\text{BLER} \uparrow$$

generally causes

$$\text{retransmission load} \uparrow$$

and

$$\text{goodput} \downarrow .$$

This connects HARQ directly to link adaptation.

A slightly more conservative MCS may reduce retransmissions enough to improve overall throughput.

11.17 Relationship Between FEC and HARQ

FEC and HARQ solve related but different problems.

FEC attempts to repair errors immediately:

$$\boxed{\text{FEC} = \text{correct errors using redundancy already transmitted.}}$$

HARQ provides additional information when the existing redundancy is not sufficient:

$$\boxed{\text{HARQ} = \text{obtain more reliability after decoding failure.}}$$

Thus the normal order is

$$\boxed{\text{receive} \rightarrow \text{FEC decode} \rightarrow \text{CRC} \rightarrow \text{HARQ if necessary.}}$$

Retransmission is therefore not the first response to a symbol error.

The receiver first attempts to correct the errors locally.

11.18 Relationship to Link Adaptation

Link adaptation and HARQ operate together.

The MCS controls how difficult the initial transmission is to decode.

HARQ recovers occasional failures.

If the MCS is too conservative,

$$\text{BLER}$$

may be nearly zero, but throughput is unnecessarily low.

If the MCS is too aggressive,

$$\text{BLER}$$

becomes large and HARQ consumes excessive airtime.

The desired operating point is therefore somewhere between these extremes.

Conceptually,

$$\boxed{\text{MCS selection} \rightarrow \text{initial reliability} \rightarrow \text{HARQ demand} \rightarrow \text{goodput.}}$$

11.19 Fundamental Perspective

CRC, ARQ, and HARQ form a hierarchy of error handling.

CRC detects whether decoding succeeded:

$$\boxed{\text{CRC} \rightarrow \text{detect failure.}}$$

ARQ responds to failure by retransmitting:

$$\boxed{\text{ARQ} \rightarrow \text{send again.}}$$

HARQ improves retransmission by preserving and combining information:

$$\boxed{\text{HARQ} \rightarrow \text{reuse previous observations.}}$$

The complete error-recovery path is therefore

$$\boxed{\text{soft demodulation} \rightarrow \text{FEC decoding} \rightarrow \text{CRC} \rightarrow \text{ACK/NACK} \rightarrow \text{HARQ if needed.}}$$

The central engineering principle is that a failed first transmission does not need to be wasted.

The receiver can preserve the information it already obtained and use later transmissions to increase confidence until the block becomes decodable.

Chapter 12

Block Error Rate, Burst Errors, and Interleaving

A wireless receiver ultimately succeeds or fails at decoding complete blocks of transmitted data.

Block Error Rate, or BLER, measures how often those blocks fail.

The distribution of errors within a block also matters. Errors may occur independently, or they may arrive in bursts because of fading, interference, or other short-duration impairments.

Interleaving helps protect coded blocks by spreading nearby coded bits across different transmission resources.

12.1 Block Error Rate

Block Error Rate is defined as

$$\boxed{\text{BLER} = \frac{N_{\text{failed blocks}}}{N_{\text{total blocks}}}.}$$

For example, suppose

$$1000$$

blocks are transmitted and

$$80$$

fail decoding.

Then

$$\text{BLER} = \frac{80}{1000} = 0.08 = 8\%.$$

A block is usually treated as failed when the receiver cannot decode it correctly and its error-detection check, such as a CRC, fails.

Thus BLER is a practical measure of how often the complete physical-layer decoding process fails.

12.2 BLER Versus BER

Bit Error Rate and Block Error Rate describe different quantities.

Bit Error Rate measures the fraction of incorrect bits:

$$\mathrm{BER} = \frac{N_{\text{incorrect bits}}}{N_{\text{total bits}}}.$$

BLER instead measures whether an entire coded block succeeds or fails.

This distinction matters because FEC operates over groups of bits.

A block may contain several unreliable raw bit observations but still be decoded correctly.

Thus,

$$\boxed{\mathrm{BER}_{\mathrm{pre-FEC}} > 0}$$

does not necessarily imply

$$\boxed{\mathrm{BLER} > 0.}$$

BLER is often more useful for link adaptation because failed blocks usually trigger retransmission.

12.3 The Cost of High BLER

A failed block does not necessarily represent permanently lost data.

HARQ or another retransmission mechanism may recover it later.

However, retransmission is not free.

A high BLER increases

- retransmission traffic,
- latency,
- scheduler load,
- power consumption,
- radio-resource usage.

The cost therefore appears as reduced useful throughput.

A simple approximation is

$$R_{\text{goodput}} \approx R_{\text{PHY}}(1 - \mathrm{BLER}).$$

This approximation ignores detailed HARQ and protocol overhead, but it illustrates the basic effect.

12.4 Example of BLER and Goodput

Suppose one MCS provides

$$500 \text{ Mbps}$$

but produces

$$40\%$$

BLER.
The approximate first-transmission goodput is

$$500(1 - 0.40) = 300 \text{ Mbps}.$$

Now suppose a more conservative MCS provides

$$420 \text{ Mbps}$$

with

$$5\%$$

BLER.
Then

$$420(1 - 0.05) = 399 \text{ Mbps}.$$

The lower nominal PHY rate therefore delivers more useful data.
This reinforces the principle

$$\boxed{\text{maximum PHY rate} \neq \text{maximum goodput}.}$$

12.5 BLER and Link Adaptation

BLER provides direct feedback about whether the selected MCS is appropriate.
If the MCS is too aggressive,

$$\text{BLER} \uparrow .$$

If the MCS is very conservative,

$$\text{BLER} \downarrow,$$

but available channel capacity may be underused.
A practical link-adaptation loop therefore uses both

$$\text{estimated channel quality}$$

and

$$\text{observed BLER}.$$

Conceptually,

$$\boxed{\text{SINR} \rightarrow \text{MCS} \rightarrow \text{BLER} \rightarrow \text{MCS adjustment.}}$$

12.6 Why Errors Often Occur in Bursts

Wireless errors are not always statistically independent.
A short impairment may corrupt many neighboring symbols.
Examples include

- a temporary deep fade,
- a short interference burst,
- blockage of a propagation path,
- a narrowband frequency fade,
- a transient synchronization problem.

A conceptual error sequence might appear as

$$XXXXXX\_\_\_\_\_,$$

where

$$X$$

represents corrupted information.
This is called a burst error.

12.7 Why Burst Errors Are Difficult

Suppose an FEC code can tolerate several uncertain bits spread throughout a codeword.
If those errors are distributed approximately as

$$X\_\_\_X\_\_\_X\_\_\_X,$$

the remaining reliable information may allow the decoder to reconstruct the block.
By contrast, a concentrated burst such as

$$XXXXX\_\_\_\_\_$$

may erase a large amount of information from one region of the codeword.
The total number of errors may be similar, but their distribution is different.
Thus,

$$\boxed{\text{error location and correlation can matter in addition to error count.}}$$

12.8 Burst Errors in OFDM

OFDM can experience burst-like errors in both time and frequency.

A short interference event may corrupt several neighboring OFDM symbols.

A frequency-selective fade may corrupt several neighboring subcarriers.

Thus the impairment may be concentrated in

time,

frequency,

or both.

For example, if several coded bits are placed on neighboring subcarriers that all fall into the same deep fade, those bits may all become unreliable simultaneously.

12.9 Interleaving

Interleaving rearranges coded bits before transmission.

Instead of transmitting coded bits in their original order,

$$c_1, c_2, c_3, c_4, c_5, c_6, \ldots,$$

the transmitter intentionally permutes them.

The receiver later reverses this permutation before FEC decoding.

The important idea is that coded bits that are adjacent in the codeword are transmitted through different physical resources.

These resources may include

- different OFDM subcarriers,
- different OFDM symbols,
- different time intervals,
- different spatial streams.

12.10 How Interleaving Helps

Suppose a short interference burst corrupts several consecutive transmitted symbols.

Without interleaving, the corresponding errors may also be adjacent in the FEC codeword:

$$XXXXX\_\_\_\_\_.$$

With interleaving, those transmitted symbols may correspond to widely separated positions in the original codeword.

After deinterleaving, the error pattern may look more like

$$X\_\_X\_\_\_X\_\_\_X\_\_X.$$

The channel impairment has not disappeared.

Instead, its effect has been redistributed.

This is the key purpose of interleaving:

$$\boxed{\text{localized channel error} \rightarrow \text{distributed decoder error.}}$$

FEC can often handle the distributed pattern more effectively.

12.11 Interleaving Across Frequency

In OFDM, one useful strategy is to distribute coded bits across different subcarriers.

Suppose one group of neighboring subcarriers experiences a deep fade.

Without frequency interleaving, a contiguous portion of the codeword may be severely corrupted.

With interleaving, bits from one region of the codeword are spread over many frequencies.

Then a localized frequency fade affects small portions of many codeword regions rather than destroying one contiguous portion.

Thus,

$$\boxed{\text{frequency diversity} + \text{interleaving} \rightarrow \text{greater robustness to frequency-selective fading.}}$$

12.12 Interleaving Across Time

The same principle can be applied across time.

Coded bits may be distributed among several OFDM symbols or transmission intervals.

A short-duration fade or interference burst then affects only a fraction of the coded information from each region of the codeword.

This provides time diversity.

However, time interleaving introduces a tradeoff.

Spreading a codeword across a longer interval can increase latency.

Thus,

$$\boxed{\text{more time interleaving} \rightarrow \text{greater burst-error protection} \rightarrow \text{potentially greater latency.}}$$

12.13 Interleaving Does Not Correct Errors

Interleaving itself is not an error-correcting code.

It does not create additional redundancy.

Instead, it changes where coded bits are transmitted.

Thus the roles are different:

$$\boxed{\text{FEC} \rightarrow \text{provides redundancy for error correction}}$$

while

$$\boxed{\text{interleaving} \rightarrow \text{arranges errors into a pattern that FEC can handle more effectively.}}$$

The two mechanisms are therefore complementary.

12.14 The Complete Error-Handling Chain

The relationship among the mechanisms discussed so far can be summarized as

$$\boxed{\text{information bits} \rightarrow \text{FEC encoding} \rightarrow \text{interleaving} \rightarrow \text{modulation} \rightarrow \text{wireless channel.}}$$

At the receiver,

$$\boxed{\text{demodulation} \rightarrow \text{deinterleaving} \rightarrow \text{FEC decoding} \rightarrow \text{CRC check.}}$$

If the block still fails,

$$\boxed{\text{CRC failure} \rightarrow \text{HARQ retransmission.}}$$

BLER measures how often this complete first-pass decoding process fails.

12.15 Fundamental Perspective

BLER measures the final block-level consequence of all physical-layer impairments:

$$\boxed{\text{noise} + \text{interference} + \text{fading} + \text{receiver errors} \rightarrow \text{decoding success or failure.}}$$

Burst errors explain why the spatial and temporal distribution of those impairments matters.

Interleaving reduces the damage caused by localized impairments by spreading coded information across different transmission resources.

The relationship can therefore be summarized as

$$\boxed{\text{burst impairment} \rightarrow \text{interleaving} \rightarrow \text{distributed errors} \rightarrow \text{more effective FEC decoding.}}$$

BLER then provides the practical measure of whether the complete system succeeded.

Chapter 13

The Complete Adaptive Wireless Link and Engineering Principles

The individual components of a wireless system are tightly coupled.

Beamforming, gain control, channel estimation, OFDM, SINR estimation, modulation and coding, FEC, HARQ, and link adaptation all contribute to one continuous feedback process.

A useful high-level description is

$$\boxed{\text{measure} \rightarrow \text{estimate} \rightarrow \text{adapt} \rightarrow \text{transmit} \rightarrow \text{observe} \rightarrow \text{adapt again.}}$$

The objective is not simply to maximize received power, SINR, or physical-layer rate.

The objective is to deliver useful information reliably and efficiently.

13.1 Establishing the Link

Before data transmission begins, the radios must first establish a usable communication path.

For a directional system, this may involve

$$\boxed{\text{beam search} \rightarrow \text{signal detection} \rightarrow \text{synchronization} \rightarrow \text{gain adjustment} \rightarrow \text{beam refinement.}}$$

The receiver must identify the desired signal while simultaneously determining suitable spatial, timing, frequency, and gain parameters.

Once the link is established, the system typically moves from global acquisition to local tracking.

Thus,

$$\boxed{\text{acquisition} = \text{large search under uncertainty}}$$

while

$$\boxed{\text{tracking} = \text{local adaptation around a known operating point.}}$$

13.2 Channel Estimation and Equalization

After synchronization, known pilot symbols allow the receiver to estimate the wireless channel.

For OFDM subcarrier k,

$$Y[k] = H[k]X[k] + I[k] + N[k].$$

The receiver estimates

$$\hat{H}[k]$$

and uses it to compensate for channel amplitude and phase distortion.
A simple equalizer is

$$\hat{X}[k] = \frac{Y[k]}{\hat{H}[k]}.$$

Channel estimates are also used for

- beamforming,
- phase tracking,
- SINR estimation,
- link adaptation.

A recurring practical principle is therefore

many receiver algorithms are only as reliable as the channel estimate they use.

13.3 Estimating Link Quality

The receiver evaluates the quality of the resulting signal using quantities such as

$$\text{RSSI}, \qquad \text{EVM}, \qquad \text{SINR}.$$

These quantities describe different properties.
RSSI measures received energy.
SINR compares the desired signal with interference and noise:

$$\text{SINR} = \frac{P_S}{P_I + P_N}.$$

Thus,

strong received power does not necessarily imply a good communication link.

A beam containing a strong interferer may have high RSSI but poor SINR.
The best communication beam is therefore often the beam that provides the best usable signal quality rather than the greatest total power.

13.4 OFDM and Frequency-Selective Channels

A wideband multipath channel generally has a frequency-dependent response

$$H(f).$$

OFDM divides the channel into many narrowband subcarriers, producing the approximate model

$$\boxed{Y[k] = H[k]X[k] + I[k] + N[k].}$$

The important engineering idea is

$$\boxed{\text{one difficult wideband frequency-selective channel} \rightarrow \text{many simpler narrowband channels.}}$$

The cyclic prefix protects the OFDM symbol from multipath delay, while the FFT separates the individual subcarriers.

The receiver can then estimate and equalize the channel independently across frequency.

13.5 Link Adaptation

Once channel quality has been estimated, the transmitter must decide how aggressively to use the link.

This is the role of the Modulation and Coding Scheme, or MCS.

Conceptually,

$$\boxed{\text{SINR} \rightarrow \text{MCS.}}$$

Poor channel conditions favor

$$\text{lower-order modulation} + \text{stronger coding},$$

while good channel conditions allow

$$\text{higher-order modulation} + \text{higher coding rate}.$$

Higher modulation order increases spectral efficiency but requires better signal quality.

Stronger coding improves robustness but reduces the fraction of transmitted bits carrying new information.

Thus MCS selection balances

$$\boxed{\text{throughput} \leftrightarrow \text{reliability.}}$$

13.6 Error Correction and Retransmission

A modern receiver does not immediately retransmit whenever individual symbols are corrupted.

Errors are handled in stages:

$$\boxed{\text{soft demodulation} \rightarrow \text{FEC} \rightarrow \text{CRC} \rightarrow \text{HARQ.}}$$

Soft demodulation preserves confidence information.
FEC attempts to reconstruct the transmitted data.
CRC determines whether the resulting block is valid.
If decoding still fails, HARQ provides additional information through retransmission.
A failed first transmission therefore does not necessarily represent wasted information.
The receiver can preserve soft information and combine it with later transmissions.

13.7 BLER and Goodput

The practical outcome of decoding is often measured using Block Error Rate:

$$\text{BLER} = \frac{N_{\text{failed blocks}}}{N_{\text{total blocks}}}.$$

SINR predicts how well an MCS should perform.
BLER measures how well it actually performs.
Thus,

$$\boxed{\text{SINR predicts} \rightarrow \text{BLER validates.}}$$

The highest physical-layer rate is not necessarily the best operating point.
A simple approximation is

$$R_{\text{goodput}} \approx R_{\text{PHY}}(1 - \text{BLER}).$$

An overly aggressive MCS can produce such a high BLER that useful throughput decreases.
Therefore,

$$\boxed{\text{maximum PHY rate} \neq \text{maximum goodput.}}$$

13.8 The Subsystems Are Coupled

One of the most important lessons of practical wireless design is that the individual subsystems cannot be considered independently.

Beamforming changes both desired signal power and interference:

$$\mathbf{w} \rightarrow P_S, \qquad \mathbf{w} \rightarrow P_I.$$

The resulting total power influences receiver gain:

$$P_S + P_I + P_N \rightarrow G_{\text{RX}}.$$

Gain affects ADC utilization and clipping:

$$G_{\text{RX}} \rightarrow \text{quantization and nonlinear distortion.}$$

These impairments affect channel and SINR estimation:

$$\text{receiver quality} \rightarrow \widehat{\text{SINR}}.$$

SINR affects MCS:

$$\widehat{\text{SINR}} \rightarrow \text{MCS}.$$

MCS and channel conditions determine BLER:

$$\text{MCS} + \text{channel} \rightarrow \text{BLER}.$$

BLER determines retransmission demand and therefore useful throughput:

$$\text{BLER} \rightarrow \text{HARQ} \rightarrow R_{\text{goodput}}.$$

The complete dependency chain can therefore be summarized as

$$\boxed{\text{beam} \rightarrow \text{received signal} \rightarrow \text{gain} \rightarrow \text{channel estimate} \rightarrow \text{SINR} \rightarrow \text{MCS} \rightarrow \text{BLER} \rightarrow \text{goodput}.}$$

13.9 Key Engineering Principles

The main lessons developed throughout this text can be reduced to several high-level principles.

Signal quality matters more than signal strength

RSSI tells us how much energy is received.

SINR tells us how usable the desired signal is.

Therefore,

$$\boxed{\text{high RSSI does not necessarily mean a good link.}}$$

Beamforming is spatial filtering

Beamforming changes both desired signal and interference.

The objective is therefore not simply maximum signal power, but improved post-beamforming SINR.

$$\boxed{\text{best beam} \neq \text{strongest beam}.}$$

Gain control protects information

Too little gain wastes ADC dynamic range.

Too much gain causes clipping and nonlinear distortion.

Therefore,

$$\boxed{\text{gain control balances sensitivity and linearity.}}$$

Information lost through analog saturation cannot simply be reconstructed later in DSP.

OFDM makes multipath manageable

OFDM transforms a frequency-selective wideband channel into many approximately flat subcarrier channels.

Its core model is

$$\boxed{Y[k] = H[k]X[k] + I[k] + N[k].}$$

Channel estimates drive receiver decisions

Pilots provide estimates of

$$H[k],$$

which support equalization, SINR estimation, phase tracking, and link adaptation.

Poor channel estimates therefore propagate into many later receiver decisions.

Modulation and coding trade efficiency for robustness

Higher-order modulation carries more information per symbol but requires better SINR.

Stronger coding tolerates poorer channels but consumes additional redundancy.

Together they define the MCS operating point.

SINR is not the final objective

SINR is useful because it predicts decoder performance.

But the actual chain is

$$\boxed{\text{SINR} \to \text{MCS} \to \text{BLER} \to \text{goodput.}}$$

SINR is therefore an intermediate metric.

Errors are handled in layers

The receiver uses several mechanisms before declaring data lost:

$$\boxed{\text{soft information} \to \text{FEC} \to \text{CRC} \to \text{HARQ.}}$$

Each layer provides a different form of reliability.

Adaptation is continuous

Wireless conditions change.

A practical radio must therefore continually adapt parameters such as

- beam selection,
- receiver gain,
- channel estimate,
- MCS,

- retransmission behavior.

The system is best understood as a collection of feedback loops rather than as a static signal-processing chain.

13.10 Final Perspective

The quantities studied throughout this text are not independent objectives.

Maximum RSSI is not the objective.

Maximum SINR is not the objective.

Maximum MCS is not the objective.

Minimum BLER is not the objective.

Each is an intermediate quantity used to make the wireless link operate effectively.

The ultimate engineering objective is

$$\boxed{\text{maximize useful delivered information}}$$

subject to acceptable

$$\text{reliability}, \qquad \text{latency}, \qquad \text{power}, \qquad \text{radio resources}.$$

In practical terms,

$$\boxed{\text{maximize goodput while maintaining a reliable link.}}$$

This system-level perspective connects beamforming, receiver design, OFDM, channel estimation, link adaptation, coding, and retransmission into one unified communication system.