

Wireless Communications and Optimization Notebook

Junchi Wang

January 2026

Category	Narrowband	Wideband
Key comparison	$B \ll B_c$	$B \gtrsim B_c$
Channel behavior	Frequency-flat	Frequency-selective
Multipath effect	Single effective tap	Multiple resolvable taps
ISI	No ISI	ISI present
Typical model	$y = hx + n$	$y(t) = \sum_\ell h_\ell x(t - \tau_\ell) + n(t)$
Typical systems	Single-carrier	OFDM, UWB

Table 1: Comparison between narrowband and wideband transmission

Category	Slow Fading	Fast Fading
Key comparison	$T_c \gg T_s$	$T_c \ll T_s$
Channel variation	Slow in time	Rapid in time
Cause	Low mobility	High mobility
Doppler spread	Small	Large
Channel model	Quasi-static	Time-varying
Performance metric	Outage probability	Ergodic capacity

Table 2: Comparison between slow fading and fast fading

1 The Wireless Channel

Large-scale fading represents the average signal power attenuation or path loss over large distances. It is generally slow-acting and influenced by the prominent terrain features between the transmitter and receiver.

1.1 Path Loss

Path loss is the deterministic reduction in power density as a wave propagates through space. A real-valued deterministic attenuation term modeled as $\Lambda_0 \propto R^\eta$

$$P_r = P_t K \left(\frac{R_0}{R} \right)^\eta$$

Important parameter: the path loss exponent η

- Determined empirically: $2 \leq \eta \leq 8$.

$$\text{Path Loss: } \Lambda_0|_{\text{dB}} = L_0|_{\text{dB}} + 10\eta \log_{10} \left(\frac{R}{R_0} \right)$$

- L_0 is the deterministic path-loss at a reference distance R_0

1.2 Shadowing

Shadowing occurs when large objects (buildings, hills) obstruct the main propagation path. It is modeled as a random variable with a **Log-Normal distribution**.

$$S|_{\text{dB}} \sim \mathcal{N}(0, \sigma_S^2) \quad (1)$$

is a zero-mean Gaussian random variable representing the random shadowing effect.

Small-scale fading refers to the rapid fluctuations of signal amplitude and phase over very short distances (wavelength scale), primarily caused by **multipath propagation**.

1.3 Multipath Effects

- **Rayleigh Fading:** Used when there is no dominant Line-of-Sight (LoS) path.
- **Rician Fading:** Used when a strong LoS component exists alongside multiple reflected paths.

For the Rayleigh Fading Channel, assuming that the signal reaches the receiver via a large number of paths of similar energy,

- Central Limit Theorem used to explain this model
- h is modeled such that its real and imaginary parts are i.i.d. zero mean Gaussian variables of variance σ^2 (circularly symmetric complex Gaussian variable).
- Recall $\mathcal{E}\{|h|^2\} = 2\sigma^2 = 1$.

2 Fading and Diversity

The system model can be written as,

$$y = \sqrt{E_s}hc + n \quad (2)$$

where the power of symbol $\mathbb{E}[\|c\|^2]$ equals 1 and the noise n follows $N(0, \sigma^2)$.

2.1 Maximum Likelihood Detection

- Decision rule: choose the hypothesis that maximizes the conditional density

$$\arg \max_x p(y|x) = \arg \max_x \log p(y|x)$$

- If real AWGN $y = x + n$ with $n \sim N(0, \sigma_n^2)$,

$$p(y|x) = \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(-\frac{(y-x)^2}{2\sigma_n^2}\right)$$

and

$$\arg \max_x p(y|x) = \arg \min_x (y-x)^2$$

- If $y = \sqrt{E_s}hc + n$, the ML decision rule becomes

$$\arg \min_c |y - \sqrt{E_s}hc|^2$$

2.2 array gain and diversity gain

- *Array Gain*: increase in average output SNR (i.e., at the input of the detector) relative to the single-branch average SNR ρ

$$g_a \triangleq \frac{\bar{\rho}_{out}}{\bar{\rho}} = \frac{\bar{\rho}_{out}}{\rho}$$

- *Diversity Gain*: increase in the error rate slope as a function of the SNR. Defined as the negative slope of the log-log plot of the average error probability \bar{P} versus SNR

$$g_d^o(\rho) \triangleq -\frac{\log_2(\bar{P})}{\log_2(\rho)}.$$

The diversity gain is commonly taken as the asymptotic slope, i.e., for $\rho \rightarrow \infty$.

2.3 SIMO System

The SIMO system model is,

$$\mathbf{y} = \sqrt{E_s} \mathbf{h} \mathbf{c} + \mathbf{n} \quad (3)$$

The gain combining can be described as $z = \mathbf{g}\mathbf{y}$

2.4 Maximal Ratio Combining

- *Maximal Ratio Combining*: weights are chosen as $g_n = h_n^*$.

- MRC maximizes the output SNR with white noise

$$\bar{\rho}_{out} = \mathcal{E} \left\{ \frac{E_s |\sum_{n=1}^{n_r} h_n^* h_n|^2}{\|\mathbf{h}\|^2 \sigma_n^2} \right\} = \frac{E_s}{\sigma_n^2} \mathcal{E} \left\{ \frac{\|\mathbf{h}\|^4}{\|\mathbf{h}\|^2} \right\} = \rho \mathcal{E} \left\{ \|\mathbf{h}\|^2 \right\} = \rho n_r.$$

The array gain g_a is thus always equal to n_r , or equivalently, the output SNR is the sum of the SNR levels of all branches (holds true irrespective of the correlation between the branches).

- For BPSK transmission, the symbol error rate reads as

$$\bar{P} = \int_0^\infty Q(\sqrt{2\rho u}) p_u(u) du$$

where $u = \|\mathbf{h}\|^2$ is χ^2 distribution with $2n_r$ degrees of freedom when the different channels are i.i.d. Rayleigh

$$p_u(u) = \frac{1}{(n_r - 1)!} u^{n_r - 1} e^{-u}.$$

At high SNR, \bar{P} becomes

$$\bar{P} = (4\rho)^{-n_r} \binom{2n_r - 1}{n_r}.$$

The diversity gain is again equal to n_r .

2.5 MMSE Combining

- So far noise was white Gaussian. When the noise (and interference) is colored, MRC is not optimal anymore.
- Let us denote the combined noise plus interference signal as \mathbf{n}_i such that $\mathbf{y} = \sqrt{E_s} \mathbf{h} c + \mathbf{n}_i$.
- An optimal gain combining technique is the minimum mean square error (MMSE) combining, where the weights are chosen in order to minimize the mean square error between the transmitted symbol c and the combiner output z , i.e.,

$$\mathbf{g}^* = \arg \min_{\mathbf{g}} \mathcal{E}_{\mathbf{n}_i, c} \{ |\mathbf{g}\mathbf{y} - c|^2 \} .$$

- Two popular solutions:
 - The optimal weight vector \mathbf{g}^* is given by

$$\mathbf{g}^* = \alpha \mathbf{h}^H \mathbf{R}_{\mathbf{yy}}^{-1},$$

where $\mathbf{R}_{\mathbf{yy}} = \mathcal{E} \{ \mathbf{yy}^H \}$ is the covariance matrix of the received signal. α is a scalar.

- The optimal weight vector \mathbf{g}^* is given by

$$\mathbf{g}^* = \beta \mathbf{h}^H \mathbf{R}_{\mathbf{n}_i}^{-1},$$

where $\mathbf{R}_{\mathbf{n}_i} = \mathcal{E} \{ \mathbf{n}_i \mathbf{n}_i^H \}$ is the covariance matrix of the combined noise plus interference signal \mathbf{n}_i . β is a scalar.

Those two solutions have the same direction (more later).

- The Signal to Interference plus Noise Ratio (SINR) at the output of the MMSE combiner simply writes as $\rho_{out} = E_s \mathbf{h}^H \mathbf{R}_{\mathbf{n}_i}^{-1} \mathbf{h}$.

2.6 MISO system

$$y = \sqrt{E_s} \mathbf{h} \mathbf{x} + n = \sqrt{E_s} \mathbf{h} \mathbf{w} c + n \quad (4)$$

where \mathbf{w} is the precoder. There are basically two different ways of achieving *direct transmit diversity*:

- when Tx has a *perfect channel knowledge*, beamforming can be performed to achieve both diversity and array gains,
- when Tx has a *partial or no channel knowledge of the channel*, space-time coding is used to achieve a diversity gain (but no array gain in the absence of any channel knowledge).

3 Channel Capacity

Definition

The capacity of an AWGN channel is

$$C = \log_2(1 + SNR) \quad [bits/s/H].$$

Definition

The capacity of a deterministic (time-invariant) SISO channel h is

$$C = \log_2(1 + \rho|h|^2) \quad [bits/s/H].$$

Instantaneous SNR is $\rho|h|^2$.

Definition

The outage probability $P_{out}(R)$ of a wireless channel with a target rate R is given by

$$P_{out}(R) = P(\log_2(1 + \rho|h|^2) < R).$$

4 Basics of MIMO

$$\mathbf{y} = \sqrt{E_s} \mathbf{H} \mathbf{x} + \mathbf{n} \quad (5)$$

where $\mathbf{y} \in \mathbb{C}^{n_r \times 1}$, $\mathbf{H} \in C^{n_r \times n_t}$ and $\mathbf{x} \in C^{n_t \times 1}$

- when Tx has a *perfect channel knowledge*: (dominant and multiple) eigenmode transmission
- when Tx has *no knowledge of the channel*: space-time coding (with $\mathbf{x}_k = \mathbf{c}_k$)

4.1 Dominant Eigenmodes

MIMO with Perfect Transmit Channel Knowledge

Extension of Matched Beamforming to MIMO:

$$\mathbf{y} = \sqrt{E_s} \mathbf{H} \mathbf{x} + \mathbf{n} = \sqrt{E_s} \mathbf{H} \mathbf{w} \mathbf{c} + \mathbf{n},$$

$$z = \mathbf{g} \mathbf{y} = \sqrt{E_s} \mathbf{g} \mathbf{H} \mathbf{w} \mathbf{c} + \mathbf{g} \mathbf{n}.$$

Decompose:

$$\mathbf{H} = \mathbf{U}_{\mathbf{H}} \boldsymbol{\Sigma}_{\mathbf{H}} \mathbf{V}_{\mathbf{H}}^H,$$

$$\boldsymbol{\Sigma}_{\mathbf{H}} = \text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_{r(\mathbf{H})}\}.$$

Received SNR is maximized by matched filtering, leading to:

$$\bullet \quad \mathbf{w} = \mathbf{v}_{max}$$

$$\bullet \quad \mathbf{g} = \mathbf{u}_{max}^H$$

where \mathbf{v}_{max} and \mathbf{u}_{max} are respectively the right and left singular vectors corresponding to the maximum singular value of \mathbf{H} , $\sigma_{max} = \max\{\sigma_1, \sigma_2, \dots, \sigma_{r(\mathbf{H})}\}$. Equivalent channel: $z = \sqrt{E_s} \sigma_{max} c + \tilde{n}$ where $\tilde{n} = \mathbf{g} \mathbf{n}$ has a variance equal to σ_n^2 .

4.2 Multiple Eigenmodes

Assume $n_r \geq n_t$ and that $r(\mathbf{H}) = n_t$, i.e., n_t singular values in \mathbf{H} . Spreading symbols over all non-zero eigenmodes of the channel:

- Tx side: multiply the input vector \mathbf{c} ($n_t \times 1$) using $\mathbf{V}_{\mathbf{H}}$, i.e. $\mathbf{c}' = \mathbf{V}_{\mathbf{H}} \mathbf{c}$.

- Rx side: multiply the received vector \mathbf{y} by $\mathbf{G} = \mathbf{U}_{\mathbf{H}}^{\mathbf{H}}$.
- Overall: $\mathbf{z} = \sqrt{E_s} \mathbf{G} \mathbf{H} \mathbf{x} + \mathbf{G} \mathbf{n} = \sqrt{E_s} \mathbf{U}_{\mathbf{H}}^{\mathbf{H}} \mathbf{H} \mathbf{V}_{\mathbf{H}} \mathbf{c} + \mathbf{U}_{\mathbf{H}}^{\mathbf{H}} \mathbf{n} = \sqrt{E_s} \Sigma_{\mathbf{H}} \mathbf{c} + \tilde{\mathbf{n}}$.

Channel decomposed into n_t parallel SISO channels given by $\{\sigma_1, \dots, \sigma_{n_t}\}$.

The rate achievable in the MIMO channel is the sum of the SISO channel capacities:

$$R = \sum_{k=1}^{n_t} \log_2(1 + \rho s_k \sigma_k^2),$$

where $\{s_1, \dots, s_{n_t}\}$ is the power allocation on each of the channel eigenmodes.

- The rate scales linearly in n_t .
- In general, the rate scales linearly with the rank of \mathbf{H} .

4.3 Water-Filling Algorithm

- $\Sigma_{\mathbf{H}} = \text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_{r(\mathbf{H})}\}$ and $\lambda_k \triangleq \sigma_k^2$

The power allocation strategy $\{s_1, \dots, s_n\} = \{s_1^*, \dots, s_n^*\}$ that maximizes $\sum_{k=1}^n \log_2(1 + \rho \lambda_k s_k)$ under the power constraint $\sum_{k=1}^n s_k = 1$, is given by the water-filling solution:

$$s_k^* = \left(\mu - \frac{1}{\rho \lambda_k} \right)^+, \quad k = 1, \dots, n$$

where μ is chosen so as to satisfy the power constraint $\sum_{k=1}^n s_k^* = 1$.

4.4 Partial Transmit Channel Knowledge

Proposition: In i.i.d. Rayleigh fading channels, the ergodic capacity with CDIT is achieved under an equal power allocation scheme $Q = \frac{I_{nt}}{n_t}$, i.e.,

$$\bar{C}_{CDIT} = \bar{I}_e = \mathbb{E} \left\{ \log_2 \det \left[I_{nr} + \frac{\rho}{n_t} H_w H_w^H \right] \right\} = \mathbb{E} \left\{ \sum_{k=1}^n \log_2 \left[1 + \frac{\rho}{n_t} \lambda_k \right] \right\}.$$

Encoding requires a fixed-rate code (whose rate is given by the ergodic capacity) with encoding spanning many channel realizations.

5 Transmission and Reception Strategies

Spatial Multiplexing / V-BLAST / D-BLAST

- No transmit channel knowledge but we know it is i.i.d. Rayleigh fading.
- Spatial Multiplexing (SM), also called V-BLAST, is a full rate code ($r_s = n_t$) that consists in transmitting independent data streams on each transmit antenna.
- In uncoded transmissions, we assume one symbol duration ($T = 1$) and codeword C is a symbol vector of size $n_t \times 1$.

Example:

$$C = \frac{1}{\sqrt{n_t}} [c_1 \ \dots \ c_{n_t}]^T.$$

Each element c_q is a symbol chosen from a given constellation.

5.1 ML Decoding

- With instantaneous channel realizations perfectly known at the receive side, the ML decoder computes an estimate of the transmitted codeword according to

$$\hat{C} = \arg \min_{C \in \mathcal{C}} \|y - \sqrt{E_s} H C\|^2.$$

where the minimization is performed over all possible codeword vectors $C \in \mathcal{C}$.

- What is C for SM?

$$C = \frac{1}{\sqrt{n_t}} [c_1 \ \dots \ c_{n_t}]^T.$$

- Assume all streams c_q use the same M -ary constellation of points \mathcal{X} , e.g., QPSK, 64-QAM, etc. Hence a search over M^{n_t} possible transmitted vectors is needed.
- Complexity gets prohibitive if M or the number of streams increases.

5.2 Zero-Forcing (ZF) Linear Receiver

- MIMO ZF receiver acts similarly to a ZF equalizer in frequency selective channels.
- ZF filtering effectively decouples the channel into n_t parallel channels.
 - Interference from other transmitted symbols is suppressed.

- Scalar decoding may be performed on each of these channels.
- The complexity of ZF decoding is similar to SISO ML decoding, but the inversion step is responsible for the noise enhancement (especially at low SNR).
- Assuming that a symbol vector $C = \frac{1}{\sqrt{n_t}} [c_1 \ \dots \ c_{n_t}]^T$ is transmitted, the output of the ZF filter G_{ZF} is given by

$$z = G_{ZF}y = [c_1 \ \dots \ c_{n_t}]^T + G_{ZF}n.$$

where G_{ZF} inverts the channel,

$$G_{ZF} = \sqrt{\frac{n_t}{E_s}} H^\dagger.$$

with $H^\dagger = (H^H H)^{-1} H^H$ denoting the Moore-Penrose pseudo inverse.

- Then each stream can be decoded $\hat{c}_q = \arg \min_{c_q \in \mathcal{X}} |z_q - c_q|, \forall q.$

5.3 Minimum Mean Squared Error (MMSE) Linear Receiver

- Recall the other MMSE solution based on

$$R_{yy} = \sigma_n^2 I_{nr} + \sum_p \frac{E_s}{n_t} h_p h_p^H$$

and the MMSE combiner for stream q is given by

$$\begin{aligned} g_{MMSE,q} &= \sqrt{\frac{E_s}{n_t}} h_q^H \left(\sigma_n^2 I_{nr} + \sum_p \frac{E_s}{n_t} h_p h_p^H \right)^{-1} \\ &= \sqrt{\frac{E_s}{n_t}} h_q^H \left(\sigma_n^2 I_{nr} + \frac{E_s}{n_t} H H^H \right)^{-1} \\ &= \sqrt{\frac{n_t}{E_s}} h_q^H \left(\frac{n_t}{\rho} I_{nr} + H H^H \right)^{-1}. \end{aligned}$$

- We can stack up combiners of all streams into one matrix:

$$G_{MMSE} = \sqrt{\frac{n_t}{E_s}} \left(H^H H + \frac{n_t}{\rho} I_{nt} \right)^{-1} H^H$$

which is a popular representation of the MMSE filter. Alternatively, it can be written as,

$$G_{MMSE} = \sqrt{\frac{n_t}{E_s}} H^H \left(H H^H + \frac{n_t}{\rho} I_{nr} \right)^{-1}$$

- Bridge between matched filtering at low SNR and ZF at high SNR.

5.4 Successive Interference Canceler

- Successively decode one symbol (or more generally one layer/stream) and cancel the effect of this symbol from the received signal.
- Decoding order based on the SINR of each symbol/layer: the symbol/layer with the highest SINR is decoded first at each iteration.
- SM with (ordered) SIC is generally known as V-BLAST, and ZF and MMSE V-BLAST refer to SM with respectively ZF-SIC and MMSE-SIC receivers.
- The diversity order experienced by the decoded layer is increased by one at each iteration. Therefore, the symbol/layer detected at iteration i will achieve a diversity of $n_r - n_t + i$.
- Major issue: error propagation
 - The error performance is mostly dominated by the weakest stream.
 - Non-ordered SIC: diversity order approximately $n_r - n_t + 1$.
 - Ordered SIC: performance improved by reducing the error propagation caused by the first decoded stream. The diversity order remains lower than n_r .

- Broadcast Channel (BC) - downlink
- Multiple Access Channel (MAC) - uplink
- The capacity region \mathcal{C} formulates this trade-off by expressing the set of all user rates (R_1, \dots, R_K) that are simultaneously achievable.

Definition

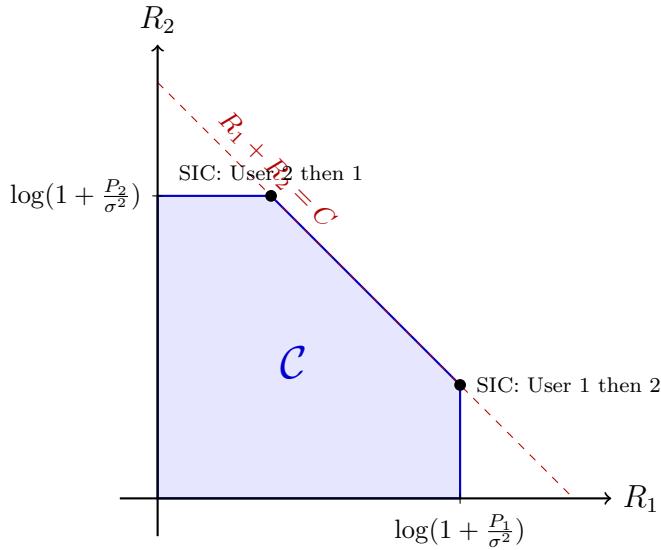
The capacity region \mathcal{C} of a channel \mathbf{H}_{ul} is the set of all rate vectors (R_1, \dots, R_K) such that simultaneously user 1 to user K can reliably communicate at rate R_1 to rate R_K , respectively.

Any rate vector not in the capacity region is not achievable (i.e. transmission at those rates will lead to errors).

Definition

The sum-rate capacity C of a capacity region \mathcal{C} is the maximum achievable sum of rates

$$C = \max_{(R_1, \dots, R_K) \in \mathcal{C}} \sum_{q=1}^K R_q.$$



5.5 Multi-user SISO

Consider the downlink case, the signal received by user q is,

$$y_q = \sqrt{E_s} h_q c + n_q \quad (6)$$

where,

$$c = \sum_q s_q c_q \quad (7)$$

Consider two-user case and the power allocations $s_1 > s_2$,

$$y_1 = h_1 s_1 c_1 + h_1 s_2 c_2 + n \quad (8)$$

$$y_2 = h_2 s_1 c_1 + h_2 s_2 c_2 + n \quad (9)$$

- Successive Interference Cancellation
- For user 1, it takes c_2 as noise and decodes c_1
- For user 2, it cancels c_1 and then decodes c_2

5.6 Multi-user MIMO

Consider the downlink case, the receive signal of user q can be written as,

$$y_q = \Lambda_q^{-1/2} \mathbf{h}_q c + n_q = \Lambda_q^{-1/2} \mathbf{h}_q \mathbf{w}_q c_q + \sum_{p \neq q} \Lambda_q^{-1/2} \mathbf{h}_q \mathbf{w}_p c_p + n_q \quad (10)$$

- includes intra-cell interference
- $h_q w_p = 0$ and $h_q w_q \neq 0$
- **ZFBF:** $HW = I$, then $W = H^\dagger$

- A *convex optimization* is one in which the objective and constraint functions are all convex², i.e.,

$$f_i(\alpha\mathbf{x} + \beta\mathbf{y}) \leq \alpha f_i(\mathbf{x}) + \beta f_i(\mathbf{y}), \forall i = 0, \dots, m \quad (11)$$

$$-\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \forall \alpha, \beta \in \mathbb{R}, \alpha + \beta = 1, \alpha \geq 0, \beta \geq 0$$

- Model Generality: LP \prec (convex) QCQP \prec SOCP \prec SDP
- Solution efficiency: LP \succ (convex) QCQP \succ SOCP \succ SDP

5.7 Linear Programming (LP)

- ▶ A set of problems with linear objective function and linear constraints¹⁰ - not so common in wireless communications.
- ▶ General LP problem

$$\min_{\mathbf{x}} \quad \mathbf{c}^T \mathbf{x} + d \quad (11a)$$

$$\text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{G}\mathbf{x} \preceq \mathbf{h} \quad (11b)$$

where $\mathbf{G} \in \mathbb{R}^{m \times n}$, $\mathbf{A} \in \mathbb{R}^{p \times n}$, \preceq refers to element-wise inequality.

- ▶ Standard LP problem

$$\min_{\mathbf{x}} \quad \mathbf{c}^T \mathbf{x} \quad (12a)$$

$$\text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \succeq \mathbf{0} \quad (12b)$$

- ▶ Inequality form LP problem

$$\min_{\mathbf{x}} \quad \mathbf{c}^T \mathbf{x} \quad (13a)$$

$$\text{s.t.} \quad \mathbf{A}\mathbf{x} \preceq \mathbf{b} \quad (13b)$$

5.8 Quadratic Constraints Quadratic programming (QCQP)

- ▶ A set of convex optimization problems with quadratic objective function and quadratic constraints¹¹

- Quadratic programming (QP) problem

$$\min_{\mathbf{x}} \quad \frac{1}{2} \mathbf{x}^T \mathbf{P} \mathbf{x} + \mathbf{q}^T \mathbf{x} + r \quad (15a)$$

$$\text{s.t.} \quad \mathbf{G} \mathbf{x} \preceq \mathbf{h}, \mathbf{A} \mathbf{x} = \mathbf{b} \quad (15b)$$

- Convex if $\mathbf{P} \in \mathbf{S}_+^n$ ($\succeq 0$)

- Minimize a convex quadratic function over a polyhedron

5.9 Second-Order Cone Programming (SOCP)

- SOCP problem¹³

$$\min_{\mathbf{x}} \quad \mathbf{f}^T \mathbf{x} \quad (20a)$$

$$\text{s.t.} \quad \|\mathbf{A}_i \mathbf{x} + \mathbf{b}_i\| \leq \mathbf{c}_i^T \mathbf{x} + d_i, \forall i = 1, \dots, m \quad (20b)$$

$$\mathbf{F} \mathbf{x} = \mathbf{g} \quad (20c)$$

- The inequality constraints $\|\mathbf{A}_i \mathbf{x} + \mathbf{b}_i\| \leq \mathbf{c}_i^T \mathbf{x} + d_i, \forall i = 1, \dots, m$ are called second-order cone (SOC) constraints
- Note constraint with $\|\cdot\|$ and not $\|\cdot\|^2$ (QCQP)
- Note $\|\mathbf{A}_i \mathbf{x} + \mathbf{b}_i\| - \mathbf{c}_i^T \mathbf{x} - d_i$ is not differentiable at 0

5.10 Semi-Definite Programming (SDP)

- A set of problems having linear objective function, linear equality constraints, and semidefiniteness constraint on variable \mathbf{X} ¹⁷
- Standard SDP problem

$$\min_{\mathbf{X}} \quad \text{Tr}(\mathbf{C} \mathbf{X}) \quad (28a)$$

$$\text{s.t.} \quad \text{Tr}(\mathbf{A}_i \mathbf{X}) = b_i, \forall i = 1, \dots, m \quad (28b)$$

$$\mathbf{X} \succeq 0 \quad (28c)$$