- We have seen how the GPM provides a computationally lightweight approach for solving the phase synchronization problem. As it turns out, the GPM can be applied to more general settings.
In this lecture, we extend the GPM for tackling a classic problem in signal processing — MIMO detection.

- Consider the linear MIMO model

$$Y = HX^* + V,$$

where $Y \in \mathbb{C}^m$ is the received signal vector, $H \in \mathbb{C}^{m \times n}$ is the channel matrix, $X^* \in \mathbb{C}^n$ is the transmitted symbol vector, and $V \in \mathbb{C}^m$ is the additive noise vector. Here, $n$ is the number of inputs and $m$ is the number of outputs.

- Typically, each symbol $X_i^*$ is drawn from some discrete constellation $\mathcal{S}$. Two common choices for $\mathcal{S}$ are

1. M-ary Phase Shift Keying (MPSK)

$$\mathcal{S} = \mathcal{S}_M = \{ \exp(2\pi ik/M) : k = 0, 1, \ldots, M-1 \}$$

2. $(4\pi^2)$-Quadrature Amplitude Modulation (QAM)

$$\mathcal{S} = \mathcal{S}_Q = \{ z \in \mathbb{C} : \text{Re}(z), \text{Im}(z) = \pm 1, \pm 3, \ldots, \pm (2u-1) \}.$$ 

Also, the entries of $V$ are assumed to be iid Gaussian RVs with mean zero and variance $\sigma_V^2$.

- Given $Y$ and $H$, the goal of MIMO detection is to recover $X^*$.
Towards that end, a natural formulation is the following maximum likelihood (ML) estimation problem:

$$\hat{X} = \arg\min_{X \in \mathcal{S}^n} \| Y - HX \|_2^2. \quad \text{(P)}$$
- Previously, a popular approach to tackling (P) is the Semidefinite relaxation Technique. A computationally less expensive alternative is the GPM, which when applied to (P) takes the following form:

\[
\begin{align*}
\tilde{W}^k &\leftarrow 2H^*(Hx^k - y) \\
x^{k+1} &\leftarrow \Pi_{\mathcal{S}^n}(x^k - \frac{2k}{m}\tilde{W}^k).
\end{align*}
\]

The efficiency of the GPM clearly depends on how efficiently the projection \(\Pi_{\mathcal{S}^n}\) can be computed. Curiously, it can be easily verified that for \(\mathcal{S} = \mathcal{S}^m\) or \(\mathcal{S} = \mathcal{C}^n\), the projection \(\Pi_{\mathcal{S}^n}\) can be efficiently computed.

- Clearly, it does not make sense to run GPM for more than \(10^n\) iterations (otherwise, we may just as well do an exhaustive search). One way to terminate the GPM is if \(x^{k+1} = x^k\) for some \(k \leq k + 1\). Such a stopping criterion will eventually be met, as the number of feasible solutions is finite.

- We are now interested in the convergence behavior of the GPM. Our goal is to prove the following:

**Theorem:** Let \(C = \frac{4}{\min_{s \neq s'} |s - s'|} < \infty\) (in particular, we have \(C = \frac{2}{\min(s_m, \Omega_m)}\) for \(\mathcal{S} = \mathcal{S}^m\) and \(C = 2\) for \(\mathcal{S} = \mathcal{C}^n\)). Suppose that

\[
\|x^k\|_\infty < \frac{1}{C} \quad \text{and} \quad \|I - \frac{2k}{m}HH^*\| \leq \beta < \frac{1}{4}.
\]

Then,

\[
\|x^{k+1} - x^k\|_2 \leq 4\beta \|x^k - x^k\|_2.
\]
In particular, after at most
\[ k^* = \left\lceil \log\left( \frac{2}{c \|x^0 - x^k\|_2} \right) / \log(4\beta) \right\rceil \]
iterations, we have \( x^k = x^* \) for all \( k > k^* \); i.e., the GPM admits finite convergence.

**Proof:** By definition of \( y \) and the GPM iterations, we have
\[
x^k = x^k - \frac{\alpha_k}{m} y^k
= x^k + (I - \frac{2\alpha_k}{m} H^H)(x^k - x^*) + \frac{2\alpha_k}{m} H^H y
\]
and \( x^{k+1} = T_{h^n}(z^k) \). Let \( w^k = (I - \frac{2\alpha_k}{m} H^H)(x^k - x^*) \) and \( J_k = \{ j : |w_{kj}^k| > \frac{1}{c} \} \). Then, by assumption of the theorem, we have
\[
|z_{lj}^k - x_{lj}^*| < \frac{\epsilon}{c} \quad \text{for} \quad l \notin J_k.
\]
This yields \( x_{lj}^{k+1} = T_{h^n}(z_{lj}^k) = x_{lj}^* \). To proceed, we need the following result:

**Proposition:** Let \( z \in \mathbb{C}^n \) and \( x \in \mathbb{R}^n \) be given, where \( S = S_m \) or \( S = S_n \). Then, we have
\[
\| T_{h^n}(z) - x \|_2 \leq 2 \| z - x \|_2.
\]

Assuming the Proposition, we compute
\[
\| x^{k+1} - x^* \|_2 = \| x_{J_k}^{k+1} - x_{J_k}^* \|_2
\leq 2 \| z_{J_k}^k - x_{J_k}^* \|_2
\leq 2 \| w_{J_k}^k + (\frac{2\alpha_k}{m} H^H y_{J_k}) \|_2
\leq 4 \| w_{J_k}^k \|_2 \quad \text{(since} \quad \| \frac{2\alpha_k}{m} H^H y_{J_k} \|_\infty < \frac{\epsilon}{c} \leq |w_{J_k}^k| \text{for} \quad j \in J_k \text{)}
\leq 4 \| w_{J_k}^k \|_2
\leq 4 \beta \| x^k - x^* \|_2 \quad \text{(since} \quad \| I - \frac{2\alpha_k}{m} H^H H \| \leq \beta \).
Lastly, note that when $\|x^{k_m} - x^*\|_2 < \frac{1}{2}$, then we must have $x^{k_m} = x^*$. This yields the bound on $k^*).

Proof of Proposition We prove the inequality for the case where $\delta = \delta_m$ and leave the case $\delta = \delta_n$ to the reader. Recall that we have previously shown that

$$\min_{r > 0} |r e^{i\phi} - 1|^2 = \begin{cases} 1 & \text{if } \phi \in \left[\frac{\pi}{2}, \frac{3\pi}{2}\right] \\ \sin^2 \phi & \text{if } \phi \in \left[0, \frac{\pi}{2}\right) \cup \left(\frac{3\pi}{2}, 2\pi\right). \end{cases}$$

Hence, if $z = r e^{i\phi}$ is such that $\phi \in \left[\frac{\pi}{2}, \frac{3\pi}{2}\right]$, then

$$|\Pi_{\delta_m}(z) - 1| \leq 2 \leq 2 |r e^{i\phi} - 1| \text{ for any } r > 0.$$ 

Otherwise, for $\phi \in \left(0, \frac{\pi}{2}\right)$, if $\phi \geq \arg(\Pi_{\delta_m}(z))$, then

$$|\Pi_{\delta_m}(z) - 1| \leq |e^{i\phi} - 1| \leq 2 |\sin \phi| \leq 2 |r e^{i\phi} - 1| \text{ for any } r > 0.$$ 

Else, we have $\phi < \arg(\Pi_{\delta_m}(z))$. In this case, we must have $\phi \geq \frac{1}{2} \arg(\Pi_{\delta_m}(z))$. Hence,

$$|\Pi_{\delta_m}(z) - 1| = 2 |\sin \frac{\arg(\Pi_{\delta_m}(z))}{2}| \leq 2 |\sin \phi| \leq 2 |r e^{i\phi} - 1| \text{ for any } r > 0.$$ 

A similar argument establishes the inequality for $\phi \in \left(\frac{3\pi}{2}, 2\pi\right).$

As it turns out, the condition in the Theorem 1 can be satisfied with high probability if the entries of $H$ are iid complex Gaussian RVs.

Theorem 2: Suppose that the entries of $H$ are iid standard complex Gaussian RVs, the entries of $v$ have variance $\sigma_v^2 \leq \frac{m}{4e \log n}$, and the aspect ratio satisfies $\gamma \leq \frac{m}{n} > \frac{20}{\beta^2} > 1$. Then, for the constant step size $\alpha_k = \frac{1}{2}$, the conditions in Theorem 1 hold with probability at least

$$1 - \frac{\log n}{m} - 4e^{-m/8} - 2e^{-n}.$$