**Optimal control and Convex optimization** **Spring 2020**

# Lecture 5: Distributed Control, QI, and SI

*Author: Yang Zheng*

**Disclaimer**: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They were developed when the author was a postdoc in Prof. Na Li's group at Harvard. Any typos should be sent to* `zhengy@eng.ucsd.edu`*.*

**Learning goals:**

1. Distributed control in static state feedback and output dynamic feedback;
2. Quadratic Invariance (QI) for sparsity constraints and delay constraints;
3. QI in finite horizon, and gradient dominance;
4. Sparsity Invariance (SI) in both static feedback and dynamic feedback;

## 1 Recap

The problem setup is as follows: we consider continuous-time linear time-invariant (LTI) systems of the form

$$\begin{aligned} \dot{x} &= Ax + B_1 w + B_2 u, \\ z &= C_1 x + D_{11} w + D_{12} u, \\ y &= C_2 x + D_{21} w + D_{22} u, \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^n, u \in \mathbb{R}^m, w \in \mathbb{R}^d, y \in \mathbb{R}^p, z \in \mathbb{R}^q$ are the state vector, control action, external disturbance, measurement, and regulated output, respectively. Consider a dynamic output feedback controller $\mathbf{u} = \mathbf{K}\mathbf{y}$, where $\mathbf{K}$ has a state-space realization

$$\begin{aligned} \dot{\xi} &= A_k \xi + B_k y, \\ u &= C_k \xi + D_k y, \end{aligned} \tag{2}$$

where $\xi \in \mathbb{R}^{n_k}$ is the internal state of controller $\mathbf{K}$.

We have introduced the following optimal control problem

$$\begin{aligned} \min_{\mathbf{K}} \quad & \|\mathbf{P}_{11} + \mathbf{P}_{12}\mathbf{K}(I - \mathbf{P}_{22}\mathbf{K})^{-1}\mathbf{P}_{21}\| \\ \text{subject to} \quad & \mathbf{K} \in \mathcal{C}_{\text{stab}}, \end{aligned} \tag{3}$$

and its corresponding state-space version (where we have assumed that $D_{22} = 0$) is

$$\begin{aligned} \min_{A_k, B_k, C_k, D_k} \quad & \left\| \left[ \begin{array}{cc|c} A + B_2 D_k C_2 & B_2 C_k & B_1 + B_2 D_k D_{21} \\ B_k C_2 & A_k & B_k D_{21} \\ \hline C_1 + D_{12} D_k C_2 & D_{12} C_k & D_{11} + D_{12} D_k D_{21} \end{array} \right] \right\| \\ \text{subject to} \quad & \begin{bmatrix} A + B_2 D_k C_2 & B_2 C_k \\ B_k C_2 & A_k \end{bmatrix} \text{ is stable.} \end{aligned} \tag{4}$$

Both (3) and (4) are non-convex in its present form. Note that the formulation (3) or (4) is very general, including LQR/LQG/$\mathcal{H}_2$/$\mathcal{H}_\infty$ optimal control [8].

We have shown that (3) or (4) is equivalent to

- An convex problem in Youla parameterization:

$$
\begin{aligned}
\min_{\mathbf{Q}} \quad & \|\mathbf{T}_{11} + \mathbf{T}_{12}\mathbf{Q}\mathbf{T}_{21}\| \\
\text{subject to} \quad & \mathbf{Q} \in \mathcal{RH}_\infty.
\end{aligned} \tag{5}
$$

  where $\mathbf{T}_{11} = \mathbf{P}_{11} + \mathbf{P}_{12}\mathbf{V}_r\mathbf{M}_l\mathbf{P}_{21}, \mathbf{T}_{12} = -\mathbf{P}_{12}\mathbf{M}_r$, and $\mathbf{T}_{21} = \mathbf{M}_l\mathbf{P}_{21}$. The controller is recovered by $\mathbf{K} = (\mathbf{V}_r - \mathbf{M}_r\mathbf{Q})(\mathbf{U}_r - \mathbf{N}_r\mathbf{Q})^{-1}$.

- An convex problem in the input-output parameterization

$$
\begin{aligned}
\min_{\mathbf{Y},\mathbf{U},\mathbf{W},\mathbf{Z}} \quad & \|\mathbf{P}_{11} + \mathbf{P}_{12}\mathbf{U}\mathbf{P}_{21}\| \\
\text{subject to} \quad & \begin{bmatrix} I & -\mathbf{P}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{Y} & \mathbf{W} \\ \mathbf{U} & \mathbf{Z} \end{bmatrix} = \begin{bmatrix} I & 0 \end{bmatrix}, \\
& \begin{bmatrix} \mathbf{Y} & \mathbf{W} \\ \mathbf{U} & \mathbf{Z} \end{bmatrix} \begin{bmatrix} -\mathbf{P}_{22} \\ I \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix}, \\
& \mathbf{Y}, \mathbf{U}, \mathbf{W}, \mathbf{Z} \in \mathcal{RH}_\infty.
\end{aligned} \tag{6}
$$

  The optimal controller is recovered by $\mathbf{K} = \mathbf{U}\mathbf{Y}^{-1}$.

- An convex problem in the system-level parameterization

$$
\begin{aligned}
\min_{\mathbf{R},\mathbf{M},\mathbf{N},\mathbf{L}} \quad & \left\| \begin{bmatrix} C_1 & D_{12} \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{N} \\ \mathbf{M} & \mathbf{L} \end{bmatrix} \begin{bmatrix} B_1 \\ D_{21} \end{bmatrix} + D_{11} \right\| \\
\text{subject to} \quad & \begin{bmatrix} sI - A & -B_2 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{N} \\ \mathbf{M} & \mathbf{L} \end{bmatrix} = \begin{bmatrix} I & 0 \end{bmatrix}, \\
& \begin{bmatrix} \mathbf{R} & \mathbf{N} \\ \mathbf{M} & \mathbf{L} \end{bmatrix} \begin{bmatrix} sI - A \\ -C_2 \end{bmatrix} = \begin{bmatrix} I \\ 0 \end{bmatrix}, \\
& \mathbf{R}, \mathbf{M}, \mathbf{N} \in \mathcal{RH}_\infty, \quad \mathbf{L} \in \mathcal{RH}_\infty.
\end{aligned} \tag{7}
$$

  The controller is recovered by $\mathbf{K} = \mathbf{L} - \mathbf{M}\mathbf{R}^{-1}\mathbf{N}$.

In addition, we have shown that the state-feedback $\mathcal{H}_2$ optimal control is equivalent to

$$
\begin{aligned}
\min_{P,X,Z} \quad & \operatorname{trace}(Z) \\
\text{subject to} \quad & (AP + B_2X) + (AP + B_2X)^\mathsf{T} + B_1B_1^\mathsf{T} \prec 0, \\
& \begin{bmatrix} Z & C_1P + D_{12}X \\ (C_1P + D_{12}X)^\mathsf{T} & P \end{bmatrix} \succ 0,
\end{aligned} \tag{8}
$$

and the optimal $\mathcal{H}_2$ optimal state feedback gain is recovered by $D_k = XP^{-1}$.

In this lecture, we focus on the classical distributed control problem, where a subspace constraint is imposed on the controller $\mathbf{K}$. It is known that the notion of Quadratic Invariance [6] allows deriving an equivalent convex formulation. More recently, a notion of Sparsity Invariance [3] generalizes the QI and allows to derive convex restriction of the largest class of distributed control problems with sparsity constraints. We also talk about Youla parameterization in the finite horizon, and prove a gradient dominance condition for distributed control problems with QI constraints [1].

# 2 Classical distributed control and Quadratic Invariance

A canonical problem one would like to solve in distributed control is to minimize a norm of the closed-loop map subject to a subspace constraint as follows

$$\begin{aligned} \min_{\mathbf{K}} \quad & \|\mathbf{P}_{11} + \mathbf{P}_{12}\mathbf{K}(I - \mathbf{P}_{22}\mathbf{K})^{-1}\mathbf{P}_{21}\| \\ \text{subject to} \quad & \mathbf{K} \in \mathcal{C}_{\text{stab}}, \\ & \mathbf{K} \in \mathcal{S}, \end{aligned} \tag{9}$$

where $\mathcal{S}$ is a subspace demoting sparsity or delay constraints on the controller. After applying the change of variables in Youla, input-output, or system-level parameterization, we need to introduce the following non-convex constraint on the decision variables

$$\begin{aligned} (\mathbf{V}_r - \mathbf{M}_r\mathbf{Q})(\mathbf{U}_r - \mathbf{N}_r\mathbf{Q})^{-1} \in \mathcal{S}, \\ \mathbf{U}\mathbf{Y}^{-1} \in \mathcal{S}, \\ \mathbf{L} - \mathbf{M}\mathbf{R}^{-1}\mathbf{N} \in \mathcal{S}. \end{aligned}$$

In this section, we introduce the notion of Quadratic Invariance [4] to deal with the constraints above. When $\mathcal{S}$ represents a sparsity pattern, we will introduce a generalize notion of Sparsity Invariance [3] in the next section.

## 2.1 Quadratic Invariance

**Definition 1** (Quadratic Invariance (QI) [6]). *Given a plant $\mathbf{P}_{22}$ and a subspace $\mathcal{S}$. The subspace $\mathcal{S}$ is called **quadratically invariant** under $\mathbf{P}_{22}$ if*

$$\mathbf{K}\mathbf{P}_{22}\mathbf{K} \in \mathcal{S}, \qquad \forall \mathbf{K} \in \mathcal{S}.$$

The subspace $\mathcal{S}$ can be used to represent a sparsity pattern or a delay pattern. Note that QI is an algebraic condition, which is independent of how to parameterize the set of stabilizing controllers. It is not suprising that QI can be combined with either Youla parameterization, SLP or IOP. In particular, under the notion of QI, we have the following result:

**Theorem 1** (QI with the IOP). *If $\mathcal{S}$ is QI under $\mathbf{P}_{22}$, then*

1. *We have*

$$\mathcal{C}_{stab} \cap \mathcal{S} = \{\mathbf{K} = \mathbf{U}\mathbf{Y}^{-1} \mid \mathbf{Y}, \mathbf{U}, \mathbf{W}, \mathbf{Z} \textit{ are in the affine subspace of } (6), \mathbf{U} \in \mathcal{S}\}. \tag{10}$$

2. *Problem* (9) *can be equivalently formulated as a convex problem*

$$\begin{aligned} \min_{\mathbf{Y},\mathbf{U},\mathbf{W},\mathbf{Z}} \quad & \|\mathbf{P}_{11} + \mathbf{P}_{12}\mathbf{U}\mathbf{P}_{21}\| \\ \textit{subject to} \quad & \textit{affine constraint in } (6), \\ & \mathbf{U} \in \mathcal{S}. \end{aligned} \tag{11}$$

*Proof.* It is easy to see that the second point directly follows the result in point 1. Here we prove the equivalence in (10).

$\Rightarrow$: Given a controller $\mathbf{K} \in \mathcal{C}_{\text{stab}}$, we have already proved that there exist $\mathbf{Y}, \mathbf{U}, \mathbf{W}, \mathbf{Z}$ satisfying the affine constraints in (6) and $\mathbf{K} = \mathbf{U}\mathbf{Y}^{-1}$. Considering the affine constraint in (6), we have

$$\mathbf{Y} - \mathbf{P}_{22}\mathbf{U} = I \quad \Rightarrow \quad \mathbf{Y} = I + \mathbf{P}_{22}\mathbf{U}.$$

Then, we have

$$\mathbf{K} = \mathbf{U}(I + \mathbf{P}_{22}\mathbf{U})^{-1} \quad \Rightarrow \quad \mathbf{U} = \mathbf{K}(I - \mathbf{P}_{22}\mathbf{K})^{-1}.$$

According to the Cayley-Hamilton Theorem, we have

$$(I - \mathbf{P}_{22}\mathbf{K})^{-1} = \alpha_0 + \alpha_1(I - \mathbf{P}_{22}\mathbf{K}) + \ldots + \alpha_{m-1}(I - \mathbf{P}_{22}\mathbf{K})^{m-1},$$

for some transfer functions $\alpha_k, k = 1, \ldots, m-1$. Considering the QI condition, we have

$$\mathbf{K}(I - \mathbf{P}_{22}\mathbf{K}) \in \mathcal{S}, \quad \mathbf{K}(I - \mathbf{P}_{22}\mathbf{K})^k \in \mathcal{S}, \quad k = 0, \ldots, m-1.$$

Thus, $\mathbf{U} = \mathbf{K}(I - \mathbf{P}_{22}\mathbf{K})^{-1} \in \mathcal{S}$.

$\Leftarrow$: We have proved that given any $\mathbf{Y}, \mathbf{U}, \mathbf{W}, \mathbf{Z}$ in the affine subspace of (6), the controller $\mathbf{K} = \mathbf{U}\mathbf{Y}^{-1} \in \mathcal{C}_{\text{stab}}$.

Now, it is easy to see that $\mathbf{K} = \mathbf{U}(I + \mathbf{P}_{22}\mathbf{U})^{-1}$. If $\mathbf{U} \in \mathcal{S}$ and $\mathcal{S}$ is QI with repsect to $\mathbf{P}_{22}$, it is similar to derive

$$\mathbf{K} = \mathbf{U}(I + \mathbf{P}_{22}\mathbf{U})^{-1} \in \mathcal{S}.$$

This completes the proof. □

The result in Theorem 1 shows that under the QI condition, the constraint on the controller $\mathbf{K}$ can be equivalently translated to the decision variable $\mathbf{U}$. Considering the equivalence among Youla, SLP and IOP [7], we have the following results.

**Corollary 1** (QI with the SLP)**.** *If $\mathcal{S}$ is QI under $\mathbf{P}_{22}$, then*

1. *We have*

$$\mathcal{C}_{stab} \cap \mathcal{S} = \{\mathbf{K} = \mathbf{L} - \mathbf{M}\mathbf{R}^{-1}\mathbf{N} \mid \mathbf{R}, \mathbf{M}, \mathbf{N}, \mathbf{L} \textit{ are in the affine subspace } (12), \mathbf{L} \in \mathcal{S}\}.$$

2. *Problem* (9) *can be equivalently formulated as a convex problem*

$$\begin{aligned} \min_{\mathbf{R},\mathbf{M},\mathbf{N},\mathbf{L}} \quad & \left\| \begin{bmatrix} C_1 & D_{12} \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{N} \\ \mathbf{M} & \mathbf{L} \end{bmatrix} \begin{bmatrix} B_1 \\ D_{21} \end{bmatrix} + D_{11} \right\| \\ \textit{subject to} \quad & \textit{affine constraint } (12), \\ & \mathbf{L} \in \mathcal{S}. \end{aligned} \tag{12}$$

**Corollary 2** (QI with Youla)**.** *If $\mathcal{S}$ is QI under $\mathbf{P}_{22}$, then*

1. *We have*

$$\mathcal{C}_{stab} \cap \mathcal{S} = \{\mathbf{K} = (\mathbf{V}_r - \mathbf{M}_r\mathbf{Q})(\mathbf{U}_r - \mathbf{N}_r\mathbf{Q})^{-1} \mid (\mathbf{V}_r - \mathbf{M}_r\mathbf{Q})\mathbf{M}_l \in \mathcal{S}, \mathbf{Q} \in \mathcal{RH}_\infty\}.$$

2. *Problem* (9) *can be equivalently formulated as a convex problem*

$$\begin{aligned} \min_{\mathbf{Q}} \quad & \|\mathbf{T}_{11} + \mathbf{T}_{12}\mathbf{Q}\mathbf{T}_{21}\| \\ \textit{subject to} \quad & (\mathbf{V}_r - \mathbf{M}_r\mathbf{Q})\mathbf{M}_l \in \mathcal{S}, \\ & \mathbf{Q} \in \mathcal{RH}_\infty. \end{aligned} \tag{13}$$

There are some typical examples of QI conditions; see [6] for details. When $\mathcal{S}$ is a sparsity constraint, we may need not to know the exact dynamic $\mathbf{P}_{22}$ to check whether the QI holds.

## 2.2 Special QI cases: Delay patterns and sparsity patterns

Following [4], we discuss some typical QI constraint: including delay patterns, symmetric constraints, and sparsity patterns.

**Distributed control with delays:** Suppose there are $n$ subsystems with transmission delay $t \geq 0$, propagation delay $p \geq 0$, and computational delay $c \geq 0$. We define the following allowable set of controllers: $\mathbf{K} \in \mathcal{S}$ if and only if

$$\mathbf{K} = \begin{bmatrix} D_c H_{11} & D_{t+c} H_{12} & \dots & D_{(n-1)t+c} H_{1n} \\ D_{t+c} H_{21} & D_c H_{22} & \dots & D_{(n-2)t+c} H_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ D_{(n-1)t+c} H_{n1} & D_{(n-2)t+c} H_{n2} & \dots & D_c H_{nn} \end{bmatrix}$$

where $D$ is a delay operator on $L_{2e}$ and $H_{ij} \in \mathcal{R}_p$ are transfer functions of appropriate dimensions. The corresponding system $\mathbf{G}$ is given by

$$\mathbf{G} = \begin{bmatrix} A_{11} & D_p A_{12} & \dots & D_{(n-1)p} A_{1n} \\ D_p A_{21} & A_{22} & \dots & D_{(n-2)p} A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ D_{(n-1)p} A_{n1} & D_{(n-2)p} A_{n2} & \dots & A_{nn} \end{bmatrix}$$

for some $A_{ij} \in \mathcal{R}_{\text{sp}}$. Then, we have the following result

**Theorem 2** ([4])**.** *Suppose that $\mathbf{G}$ and $\mathcal{S}$ are defined as above, then $\mathcal{S}$ is QI with respect to $\mathbf{G}$ if and only if*

$$t \leq p + \frac{1}{n-1}.$$

The proof is based on direct verification; see [4, Thoerem 22] for details.

A nice physical interpretation of Theorem 2 is that the constraint is QI *if the controllers can communicate faster than the dynamics propagate*, *i.e.*, $t \leq p$. In this case, the optimal distributed controller may be found via convex programming. Note that this interpretation has been widely used in the literature.

**Symmetric constraints:** When the plant is symmetric, the constraint of symmetric controllers is naturally QI. In particular, we have the following result:

**Theorem 3** ([4])**.** *Suppose $\mathbb{H}^n = \{A \in \mathbb{C}^{n \times n} \mid A = A^*\}$, and $\mathcal{S} = \{\mathbf{K} \in \mathcal{R}_p \mid K(j\omega) \in \mathbb{H}^n, \forall \omega \in \mathbb{R}\}$. If $\mathbf{G} \in \mathcal{R}_p$ with $\mathbf{G}(j\omega) \in \mathbb{H}^n$, then $\mathcal{S}$ is QI with respect to $\mathbf{G}$.*

**Sparsity constraints:** Many problems in distributed control can be expressed in the form of problem 9, where $\mathcal{S}$ is the set of controllers that satisfy a specified sparsity constraint. Here, we provide a computational test for quadratic invariance when the subspace $\mathcal{S}$ is defined by sparsity constraints. We first introduce some notation.

Suppose $A^{\text{bin}} \in \{0,1\}^{m \times n}$ is a binary matrix. We define the subspace

$$\text{Sparse}(A^{\text{bin}}) = \{\mathbf{B} \in \mathcal{R}_p \mid \mathbf{B}_{ij}(\omega) = 0, \text{ for all}, i, j, \text{such that } A^{\text{bin}}_{ij} = 0, \text{ for almost all } \omega \in \mathbb{R}\}.$$

Also, if $\mathbf{B} \in \mathcal{R}_p$, let $A^{\text{bin}} = \text{Pattern}(\mathbf{B})$ be the binary matrix given by

$$A_{ij}^{\text{bin}} = \begin{cases} 0, & \text{if } \mathbf{B}_{ij}(j\omega) = 0 \text{ for almost all } \omega \in \mathbb{R}, \\ 1, & \text{otherwise.} \end{cases}$$

The following result provides a computational test for quadratic invariance when $\mathcal{S}$ is defined by sparsity constraints.

**Theorem 4** ([4])**.** *Suppose $\mathcal{S} = Sparse(K^{bin})$, and let $G^{bin} = Pattern(\mathbf{G})$. Then, the following condition are equivalent:*

1. *$\mathcal{S}$ is QI with respect to $\mathbf{G}$.*
2. *$\mathbf{KGJ} \in \mathcal{S}, \forall \mathbf{K}, \mathbf{J} \in \mathcal{S}$.*
3. *$K_{ki}^{bin} G_{ij}^{bin} K_{jl}^{bin}(1 - K_{kl}^{bin}) = 0, \forall i, l = 1, \dots, n_y$ and $j, k = 1, \dots, n_u$.*

Here, we show a negative result: perfectly decentralized control is never QI except for the trivial case where no subsystem affects any other.

**Corollary 3.** *Suppose there exists $i, j$ with $i \neq j$ such that $\mathbf{G}_{ij} \neq 0$. Suppose $K^{bin}$ is diagonal and $\mathcal{S} = Sparse(K^{bin})$. Then, $\mathcal{S}$ is not QI under $\mathbf{G}$.*

## 2.3 Non-QI cases

Consider a sparsity pattern $K^{\text{bin}}$ and a subspace $\mathcal{S} = \text{Sparse}(K^{\text{Bin}})$. Given a plant $\mathbf{G}$, if $\mathcal{S}$ is not QI with respect to $\mathbf{G}$, *i.e.*,

$$K^{\text{bin}} G^{\text{bin}} K^{\text{bin}} \nleq K^{\text{bin}}.$$

In [5], the authors proposed the closest subset and superset of $\mathcal{S}$ to make it QI

- **Closest Superset:** we aim to solve the following binary optimization problem

$$\begin{aligned} \min_{Z} \quad & \mathcal{N}(Z) \\ \text{subject to} \quad & Z G^{\text{bin}} Z \leq Z \\ & K^{\text{bin}} \leq Z, \end{aligned}$$

  where $Z$ is a binary matrix and $\mathcal{N}(Z)$ denotes the number of non-zero elements in $Z$. Even though this is nonlinear integer program, it is proved in [5] that this problem adimits a polynomial time solution as follow

$$\begin{aligned} Z_0 &= K^{\text{bin}} \\ Z_{m+1} &= Z_m + Z_m G^{\text{bin}} Z_m, \quad m \geq 0 \end{aligned}$$

  which will converge within finite iterations.

- **Closet subset:** we aim to solve the following binary optimization problem

$$\begin{aligned} \max_{Z} \quad & \mathcal{N}(Z) \\ \text{subject to} \quad & Z G^{\text{bin}} Z \leq Z \\ & Z \leq K^{\text{bin}}. \end{aligned}$$

  There is no known efficient algorithms to solve the problem above.

Here, we present a dual approach to approximate the plant dynamics $\mathbf{G}$, which can be combined with robust control to provide suboptimality guarantee.

$$\begin{aligned} \min_{\mathbf{G}_0} \quad & \|\mathbf{G}_0 - \mathbf{G}\|_\infty \\ \text{subject to} \quad & K^{\text{bin}} \cdot \text{Pattern}(\mathbf{G}_0) \cdot K^{\text{bin}} \leq K^{\text{bin}}. \end{aligned} \tag{14}$$

We note that (14) is equivalent to

$$\begin{aligned} \max_{G_0} \quad & \mathcal{N}(G_0) \\ \text{subject to} \quad & G_0 \leq \text{Pattern}(\mathbf{G}) \\ & K^{\text{bin}} G_0 K^{\text{bin}} \leq K^{\text{bin}}. \end{aligned} \tag{15}$$

Unlike the nearest QI subset approach, the constraint in (15) is linear in the decision version $G_0$, and Problem (15) admits a globally optimal solution.

# 3 Distributed control in finite horizon and Gradient Dominance

Here, we discuss a result of gradient dominance for distributed control in finite horizon with QI constraints. The presentation of this section largely follows [1].

## 3.1 Problem setup

We consider time-varying linear systems in discrete-time

$$\begin{aligned} x_{t+1} &= A_t x_t + B_t u_t + w_t \,, \\ y_t &= C_t x_t + v_t \,, \end{aligned} \tag{16}$$

where $x_t \in \mathbb{R}^n$ is the system state at time $t$ affected by process noise $w_t \sim \mathcal{D}_w$ with $x_0 = \mu_0 + \delta_0$, $\delta_0 \sim \mathcal{D}_{\delta_0}$, $y_t \in \mathbb{R}^p$ is the observed output at time $t$ affected by measurement noise $v_t \sim \mathcal{D}_v$, and $u_t \in \mathbb{R}^m$ is the control input at time $t$ to be designed. We assume that the distributions $\mathcal{D}_w, \mathcal{D}_{\delta_0}$ $\mathcal{D}_v$ are bounded, have zero mean and variances of $\Sigma_w, \Sigma_{\delta_0}, \Sigma_v \succ 0$ respectively. We consider the evolution of (16) in finite-horizon for $t = 0, \dots N$, where $N \in \mathbb{N}$. By defining the matrices

$$\mathbf{A} = \text{blkdg}(A_0, \dots, A_N), \quad \mathbf{B} = \begin{bmatrix} \text{blkdg}(B_0, \dots, B_{N-1}) \\ 0_{n \times mN} \end{bmatrix}, \quad \mathbf{C} = \text{blkdg}(C_0, \dots, C_N)\,,$$

and the vectors

$$\begin{aligned} \mathbf{x} &= \begin{bmatrix} x_0^\mathsf{T} & \dots & x_N^\mathsf{T} \end{bmatrix}^\mathsf{T}, \\ \mathbf{y} &= \begin{bmatrix} y_0^\mathsf{T} & \dots & y_N^\mathsf{T} \end{bmatrix}^\mathsf{T}, \\ \mathbf{u} &= \begin{bmatrix} u_0^\mathsf{T} & \dots & u_{N-1}^\mathsf{T} \end{bmatrix}^\mathsf{T}, \\ \mathbf{w} &= \begin{bmatrix} x_0^\mathsf{T} & w_0^\mathsf{T} & \dots & w_{N-1}^\mathsf{T} \end{bmatrix}^\mathsf{T}, \\ \mathbf{v} &= \begin{bmatrix} v_0^\mathsf{T} & \dots & v_N^\mathsf{T} \end{bmatrix}^\mathsf{T}, \end{aligned}$$

and the block-down shift matrix

$$\mathbf{Z} = \begin{bmatrix} 0_{1 \times N} & 0 \\ I_N & 0_{N \times 1} \end{bmatrix} \otimes I_n\,,$$

we can write the system (16) compactly as $\mathbf{x} = \mathbf{ZAx} + \mathbf{ZBu} + \mathbf{w}$, $\mathbf{y} = \mathbf{Cx} + \mathbf{v}$, leading to

$$\mathbf{x} = \mathbf{P}_{11}\mathbf{w} + \mathbf{P}_{12}\mathbf{u}\,, \quad \mathbf{y} = \mathbf{Cx} + \mathbf{v}, \tag{17}$$

where $\mathbf{P}_{11} = (I - \mathbf{ZA})^{-1}$ and $\mathbf{P}_{12} = (I - \mathbf{ZA})^{-1}\mathbf{ZB}$. We consider linear output-feedback policies

$$u_t = K_{t,0}y_0 + K_{t,1}y_1, + \ldots, K_{t,t}y_t, t = 0, 1, \ldots, N-1.$$

More compactly

$$\mathbf{u} = \mathbf{Ky}, \quad \mathbf{K} \in \mathcal{K}\,, \tag{18}$$

where $\mathcal{K}$ is a subspace in $\mathbb{R}^{mN \times p(N+1)}$ that 1) ensures causality of $\mathbf{K}$ by setting to 0 those entries that correspond to future outputs, 2) can enforce a time-varying spatio-temporal information structure for distributed control.

## 3.2 Distributed control in finite horizon

The distributed Linear Quadratic (LQ) optimal control problem in finite-horizon is:

$$\min_{\mathbf{K} \in \mathcal{K}} \quad J(\mathbf{K})\,, \tag{19}$$

where the cost $J(\mathbf{K})$ is defined as

$$J(\mathbf{K}) := \mathbb{E}_{\mathbf{w},\mathbf{v}} \left[ \sum_{t=0}^{N-1} \left( y_t^\top M_t y_t + u_t^\top R_t u_t \right) + y_N^\top M_N y_N \right], \tag{20}$$

and $M_t \succeq 0$ and $R_t \succ 0$ for every $t$. We denote the optimal value of problem $LQ_{\mathcal{K}}$ as $J^\star$. By rearranging (17)-(18), it can be observed that $J(\mathbf{K})$ is in general a non-convex multivariate polynomial in the entries of $\mathbf{K}$. Note that $LQ_{\mathcal{K}}$ is a constrained problem over the subspace $\mathcal{K}$; it is convenient to observe that $LQ_{\mathcal{K}}$ is actually equivalent to an unconstrained problem.

**Lemma 1.** *Let $d \in \mathbb{N}$ be the dimension of $\mathcal{K}$, and the columns of $P \in \mathbb{R}^{mpN(N+1) \times d}$ be a basis of the subspace $\{vec(\mathbf{K}) |\ \forall \mathbf{K} \in \mathcal{K}\}$. Define the function $f : \mathbb{R}^d \to \mathbb{R}$ as $f(z) := J(vec^{-1}(Pz))$. Then, $LQ_{\mathcal{K}}$ is equivalent to the* unconstrained *problem*[1]

$$\min_{z \in \mathbb{R}^d} f(z)\,. \tag{21}$$

*Proof.* Since the columns of $P$ are a basis of $\mathcal{K}$, we have 1) $\forall \mathbf{K} \in \mathcal{K}$, $\exists z \in \mathbb{R}^d$ such that $\text{vec}(\mathbf{K}) = Pz$ and 2) $\forall z \in \mathbb{R}^d$, $\text{vec}^{-1}(Pz) \in \mathcal{K}$. Hence, (21) is equivalent to $LQ_{\mathcal{K}}$. □

The function $f(z)$ is generally a non-convex multivariate polynomial in $z \in \mathbb{R}^d$ which may possess multiple local-minima, thus preventing global convergence of model-free algorithms. Fortunately, $f(z)$ admits a unique global minimum if it is *gradient dominated i.e.*,

$$\mu(f(z) - J^\star) \leq \|\nabla f(z)\|_2^2\,,\ \forall z \in \mathbb{R}^d$$

for some $\mu > 0$.

[1]Throughout this lecture, $J(\mathbf{K})$ is reserved for the LQ cost function in (20) and $f(z)$ is reserved for the equivalent cost function $f(z) := J(\text{vec}^{-1}(Pz))$.

### 3.3 QI and gradient dominance

It is well-known since the work of [4] that problem (19) can be equivalently transformed into a strongly convex program if and only if QI holds, that is

$$\mathbf{K}\mathbf{C}\mathbf{P}_{12}\mathbf{K} \in \mathcal{K}, \quad \forall \mathbf{K} \in \mathcal{K} \,. \tag{22}$$

Here, we prove a *local* gradient dominance property for 1) the class of all QI instances of (19) 2) other non-QI instances of (19).

**Theorem 5** ( [1])**.** *Let $\mathcal{K}$ be QI with respect to $\mathbf{CP}_{12}$, i.e., (22) holds. For any $\delta > 0$ and initial value $z_0 \in \mathbb{R}^d$, define the sublevel set $\mathcal{G}_{10\delta^{-1}} = \{z \in \mathbb{R}^d \mid \; f(z) - J^\star \leq 10\delta^{-1}\Delta_0\}$, where $\Delta_0 := f(z_0) - J^\star$ is the initial optimality gap. Then, the following statements hold.*

1. *$\mathcal{G}_{10\delta^{-1}}$ is compact.*
2. *$f(z)$ has a unique stationary point.*
3. *$f(z)$ admits a local gradient dominance constant $\mu_\delta > 0$ over $\mathcal{G}_{10\delta^{-1}}$, that is*

$$\mu_\delta(f(z) - J^\star) \leq \|\nabla f(z)\|_2^2 \,, \; \forall z \in \mathcal{G}_{10\delta^{-1}} \,. \tag{23}$$

The notion of QI guarantees existence of a gradient dominance constant $\mu_\delta$ which is "global" on $\mathcal{G}_{10\delta^{-1}}$, for any $\delta > 0$. By inspection of (23), for every $\delta > 0$, the only stationary point contained in $\mathcal{G}_{10\delta^{-1}}$ is the global optimum, since whenever $\nabla f(z) = 0$, we have $f(z) = J^\star$.

*Proof.* to add □

## 4 Sparsity Invariance

The QI condition put a strict requirement on the subspace $\mathcal{S}$. If the subspace $\mathcal{S}$ does not satisfy the QI condition, then none of the formulations (11), (12), (13) can be used to solve the problem (9).

In this section, we introduce a new notion of Sparsity Invariance (SI) [3] that always leads to a convex restriction of (9). This SI notion aims to deal with sparsity subspace $\mathcal{S}$ only, and can be viewed as a generalization of the QI notion; see [3] for the output feedback case, and [2] for the static feedback case.

## References

[1] Luca Furieri, Yang Zheng, and Maryam Kamgarpour. Learning the globally optimal distributed lq regulator. *arXiv preprint arXiv:1912.08774*, 2019.

[2] Luca Furieri, Yang Zheng, Antonis Papachristodoulou, and Maryam Kamgarpour. On separable quadratic lyapunov functions for convex design of distributed controllers. In *2019 18th European Control Conference (ECC)*, pages 42–49. IEEE, 2019.

[3] Luca Furieri, Yang Zheng, Antonis Papachristodoulou, and Maryam Kamgarpour. Sparsity invariance for convex design of distributed controllers. *arXiv preprint arXiv:1906.06777*, 2019.

[4] Michael Rotkowitz and Sanjay Lall. A characterization of convex problems in decentralized control. *IEEE Transactions on Automatic Control*, 51(2):274–286, 2006.

[5] Michael C Rotkowitz and Nuno C Martins. On the nearest quadratically invariant information constraint. *IEEE Transactions on Automatic Control*, 57(5):1314–1319, 2011.

[6] Michael Charles Rotkowitz. *Tractable problems in optimal decentralized control*. Stanford University, 2005.

[7] Yang Zheng, Luca Furieri, Antonis Papachristodoulou, Na Li, and Maryam Kamgarpour. On the equivalence of youla, system-level and input-output parameterizations. *arXiv preprint arXiv:1907.06256*, 2019.

[8] Kemin Zhou, John Comstock Doyle, Keith Glover, et al. *Robust and optimal control*, volume 40. Prentice hall New Jersey, 1996.