Quadratic Performance of Primal-Dual Methods for Distributed Optimization

> 55th IEEE Conference on Decision and Control Las Vegas, NV

> > John W. Simpson-Porco

Bala K. Poolla, Nima Monshizadeh, and Florian Dörfler.



ET H

Eidgenössische Technische Hochschule Zürich Swiss Federal Institute of Technology Zurich



## Our 20th Century Bulk Power System

A large-scale, nonlinear, hybrid, stochastic, distributed, cyber-physical ...



#### What kind of control is used?

#### synchronous generator



#### synchronous generator power inverters





#### synchronous generator

power inverters



#### location & distributed implementation



#### synchronous generator

power inverters



#### location & distributed implementation



scaling

#### synchronous generator

power inverters



location & distributed implementation



Central control authority is fading ... how to coordinate new actuators?

scaling

# **Optimal Distributed Frequency Regulation**

#### A very incomplete literature review



**Popular Idea:** {Grid Dynamics} ∪ {Dist. Controller} = Distributed <u>Online</u> Optimization Algorithm

(many other approaches exist)

# **Optimal Distributed Frequency Regulation**

#### A very incomplete literature review



#### **Popular Idea:** {Grid Dynamics} $\cup$ {Dist. Controller}

= Distributed **Online Optimization Algorithm** 

(many other approaches exist)

# **Optimal Distributed Frequency Regulation**

#### A very incomplete literature review



# **Popular Idea:** {Grid Dynamics} $\cup$ {Dist. Controller}

= Distributed Online Optimization Algorithm

(many other approaches exist)

$$\begin{array}{ll} \underset{x \in \mathbb{R}^{n}}{\text{minimize}} & \frac{1}{2}x^{\mathsf{T}}Qx + c^{\mathsf{T}}x\\ \text{subject to} & Sx = b \,, \end{array}$$

x ∈ ℝ<sup>n</sup> is our decision variable, Q = diag(q<sub>1</sub>,...,q<sub>n</sub>) is pos. def.
S ∈ ℝ<sup>r×n</sup> with r < n is full rank, c, b are vectors</li>

Each variable  $x_i \in \mathbb{R}$  belongs to **agent**.

The constraint matrix *S* couples the agents together.

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & \frac{1}{2} x^\mathsf{T} Q x + c^\mathsf{T} x \\ \text{subject to} & S x = b \,, \end{array}$$

x ∈ ℝ<sup>n</sup> is our decision variable, Q = diag(q<sub>1</sub>,...,q<sub>n</sub>) is pos. def.
S ∈ ℝ<sup>r×n</sup> with r < n is full rank, c, b are vectors</li>

Each variable  $x_i \in \mathbb{R}$  belongs to **agent**.

The constraint matrix *S* couples the agents together.

$$\begin{array}{ll} \underset{x \in \mathbb{R}^{n}}{\text{minimize}} & \frac{1}{2}x^{\mathsf{T}}Qx + c^{\mathsf{T}}x\\ \text{subject to} & Sx = b \,, \end{array}$$

x ∈ ℝ<sup>n</sup> is our decision variable, Q = diag(q<sub>1</sub>,...,q<sub>n</sub>) is pos. def.
S ∈ ℝ<sup>r×n</sup> with r < n is full rank, c, b are vectors</li>

Each variable  $x_i \in \mathbb{R}$  belongs to **agent**.

The constraint matrix *S* couples the agents together.

$$\begin{array}{ll} \underset{x \in \mathbb{R}^{n}}{\text{minimize}} & \frac{1}{2}x^{\mathsf{T}}Qx + c^{\mathsf{T}}x\\ \text{subject to} & Sx = b \,, \end{array}$$

x ∈ ℝ<sup>n</sup> is our decision variable, Q = diag(q<sub>1</sub>,...,q<sub>n</sub>) is pos. def.
S ∈ ℝ<sup>r×n</sup> with r < n is full rank, c, b are vectors</li>

Each variable  $x_i \in \mathbb{R}$  belongs to **agent**.

The constraint matrix *S* couples the agents together.

Key Idea: Write an algorithm which seeks the KKT points of

$$L(x,\nu) = \frac{1}{2}x^{\mathsf{T}}Qx + c^{\mathsf{T}}x + \nu^{\mathsf{T}}(Sx - b).$$

Saddle Point Algorithm [Kose '56, Arrow *et al.* '58]  

$$\tau_{x}\dot{x} = -\nabla_{x}L(x,\nu) \implies \begin{array}{l} \tau_{x}\dot{x} = -Qx - c - S^{\mathsf{T}}\nu, \\ \tau_{\nu}\dot{\nu} = \nabla_{\nu}L(x,\nu) \end{array} \qquad \Rightarrow \begin{array}{l} \tau_{\nu}\dot{x} = -Qx - c - S^{\mathsf{T}}\nu, \\ \tau_{\nu}\dot{\nu} = Sx - b, \end{array}$$

**1** Dynamics are **distributed** with sparsity pattern of *S* 

- 2 Converges to unique primal-dual optimizer  $(x^*, \nu^*)$  [Feijer '10, Cherukuri '15]
- 3 In control though, stability is just the first step...

Key Idea: Write an algorithm which seeks the KKT points of

$$L(x,\nu) = \frac{1}{2}x^{\mathsf{T}}Qx + c^{\mathsf{T}}x + \nu^{\mathsf{T}}(Sx - b).$$

Opposite the second second

- 2 Converges to unique primal-dual optimizer  $(x^*, 
  u^*)$  [Feijer '10, Cherukuri '15]
- 3 In control though, stability is just the first step...

Key Idea: Write an algorithm which seeks the KKT points of

$$L(x,\nu) = \frac{1}{2}x^{\mathsf{T}}Qx + c^{\mathsf{T}}x + \nu^{\mathsf{T}}(Sx - b).$$

Saddle Point Algorithm [Kose '56, Arrow et al. '58]

$$\begin{aligned} \tau_{x}\dot{x} &= -\nabla_{x}\mathcal{L}(x,\nu) \\ \tau_{\nu}\dot{\nu} &= \nabla_{\nu}\mathcal{L}(x,\nu) \end{aligned} \implies \begin{array}{c} \tau_{x}\dot{x} &= -Qx - c - S^{\mathsf{T}}\nu \,, \\ \tau_{\nu}\dot{\nu} &= Sx - b \,, \end{aligned}$$

**1** Dynamics are **distributed** with sparsity pattern of *S* 

- ② Converges to unique primal-dual optimizer  $(x^*,
  u^*)$  [Feijer '10, Cherukuri '15]
- In control though, stability is just the first step...

Key Idea: Write an algorithm which seeks the KKT points of

$$L(x,\nu) = \frac{1}{2}x^{\mathsf{T}}Qx + c^{\mathsf{T}}x + \nu^{\mathsf{T}}(Sx - b).$$

- **1** Dynamics are **distributed** with sparsity pattern of *S*
- 2 Converges to unique primal-dual optimizer  $(x^*, \nu^*)$  [Feijer '10, Cherukuri '15]
- In control though, stability is just the first step...

Key Idea: Write an algorithm which seeks the KKT points of

$$L(x,\nu) = \frac{1}{2}x^{\mathsf{T}}Qx + c^{\mathsf{T}}x + \nu^{\mathsf{T}}(Sx - b).$$

Saddle Point Algorithm [Kose '56, Arrow *et al.* '58]  $\tau_{x}\dot{x} = -\nabla_{x}L(x,\nu) \implies \begin{array}{l} \tau_{x}\dot{x} = -Qx - c - S^{\mathsf{T}}\nu, \\ \tau_{\nu}\dot{\nu} = \nabla_{\nu}L(x,\nu) \end{array} \implies \begin{array}{l} \tau_{\nu}\dot{\nu} = Sx - b, \end{array}$ 

- **1** Dynamics are **distributed** with sparsity pattern of *S*
- 2 Converges to unique primal-dual optimizer  $(x^*, \nu^*)$  [Feijer '10, Cherukuri '15]
- In control though, stability is just the first step...

Key Idea: Write an algorithm which seeks the KKT points of

$$L(x,\nu) = \frac{1}{2}x^{\mathsf{T}}Qx + c^{\mathsf{T}}x + \nu^{\mathsf{T}}(Sx - b).$$

Saddle Point Algorithm [Kose '56, Arrow *et al.* '58]  $\tau_{x}\dot{x} = -\nabla_{x}L(x,\nu) \implies \begin{array}{l} \tau_{x}\dot{x} = -Qx - c - S^{\mathsf{T}}\nu, \\ \tau_{\nu}\dot{\nu} = \nabla_{\nu}L(x,\nu) \end{array} \implies \begin{array}{l} \tau_{\nu}\dot{\nu} = Sx - b, \end{array}$ 

- **1** Dynamics are **distributed** with sparsity pattern of *S*
- 2 Converges to unique primal-dual optimizer  $(x^*, \nu^*)$  [Feijer '10, Cherukuri '15]
- In control though, stability is just the first step...

Primal-Dual System  

$$G:\begin{cases} \tau_x \dot{x} = -Qx - c - S^{\mathsf{T}}\nu + d_c \\ \tau_\nu \dot{\nu} = Sx - b + d_b \\ z = \frac{1}{\sqrt{2}}Q^{\frac{1}{2}}x \end{cases}$$

**Note:**  $||z||_2^2 = \frac{1}{2}x^T Qx \sim Cost Function$ 

G defines an I/O mapping  $G: (d_c, d_b) \rightarrow z$ .

Primal-Dual System  

$$G:\begin{cases} \tau_x \dot{x} = -Qx - c - S^{\mathsf{T}}\nu + d_c \\ \tau_\nu \dot{\nu} = Sx - b + d_b \\ z = \frac{1}{\sqrt{2}}Q^{\frac{1}{2}}x \end{cases}$$

**Note:**  $||z||_2^2 = \frac{1}{2}x^T Qx \sim \text{Cost Function}$ 

G defines an I/O mapping  $G: (d_c, d_b) \rightarrow z$ .

Primal-Dual System  

$$G:\begin{cases} \tau_x \dot{x} = -Qx - c - S^{\mathsf{T}}\nu + d_c \\ \tau_\nu \dot{\nu} = Sx - b + d_b \\ z = \frac{1}{\sqrt{2}}Q^{\frac{1}{2}}x \end{cases}$$

**Note:**  $||z||_2^2 = \frac{1}{2}x^T Qx \sim \text{Cost Function}$ 

G defines an I/O mapping  $G: (d_c, d_b) \rightarrow z$ .

Primal-Dual System  

$$G:\begin{cases} \tau_x \dot{x} = -Qx - c - S^{\mathsf{T}}\nu + d_c \\ \tau_\nu \dot{\nu} = Sx - b + d_b \\ z = \frac{1}{\sqrt{2}}Q^{\frac{1}{2}}x \end{cases}$$

Note:  $||z||_2^2 = \frac{1}{2}x^T Qx \sim Cost$  Function

G defines an I/O mapping  $G: (d_c, d_b) \rightarrow z$ .

## Block Diagram of Primal-Dual I/O System



## Quick Review: The $\mathcal{H}_2$ System Norm

Exp. Stable Linear System  

$$\dot{x} = Ax + Bd$$
 $z = Cx$ 
 $G(s) = C(sI - A)^{-1}B$ 

$$\|G\|_{\mathcal{H}_2}^2 \triangleq \frac{1}{2\pi} \int_{\mathbb{R}} \operatorname{Tr} \left[ G^{\mathsf{T}}(-j\omega) G(j\omega) \right] \, \mathrm{d}\omega$$

Useful Interpretations:

- (i) **Steady-state output variance**  $\lim_{t\to\infty} \mathbb{E}[z(t)^{\mathsf{T}}z(t)]$  when *d* noise
- (ii) "Average" gain over all frequencies from  $d(\cdot)$  to  $z(\cdot)$

If 
$$(A, C)$$
 observable  
 $A^{\mathsf{T}}Y + YA + C^{\mathsf{T}}C = 0$ 
 $\Rightarrow$ 
 $\|G\|_{\mathcal{H}_2}^2 = \operatorname{Tr}(B^{\mathsf{T}}YB)$ 

Primal-Dual System  $G: \begin{cases} \tau_x \dot{x} = -Qx - c - S^{\mathsf{T}} \nu + d_c \\ \tau_\nu \dot{\nu} = Sx - b + d_b \\ z = \frac{1}{\sqrt{2}} Q^{\frac{1}{2}} x \end{cases} \tag{(\star)}$ 

**Theorem**: Primal-Dual Performance The  $\mathcal{H}_2$  norm of the system (\*) is

$$\|G\|_{\mathcal{H}_2}^2 = \frac{1}{4} \sum_{i=1}^n \frac{1}{\tau_{\mathsf{x},ii}} + \frac{1}{4} \sum_{k=1}^r \frac{1}{\tau_{\nu,kk}}$$

For uniform time-constants

$$||G||_{\mathcal{H}_2}^2 = \frac{1}{4\tau_x}n + \frac{1}{4\tau_\nu}r.$$



#### **Key Insights**

- **O** Norm **independent** of Q, S
- 2 Speed/Performance Trade-off
- **1** Norm **grows** with # agents



Key Insights
Norm independent of Q, S
Speed/Performance Trade-off
Norm grows with # agents



Key Insights
Norm independent of Q, S
Speed/Performance Trade-off
Norm grows with # agents



#### **Key Insights**

- **()** Norm **independent** of Q, S
- 2 Speed/Performance Trade-off
- On the second second



#### **Key Insights**

- **()** Norm **independent** of Q, S
- 2 Speed/Performance Trade-off
- 3 Norm grows with # agents



#### **Key Insights**

- **O** Norm **independent** of Q, S
- 2 Speed/Performance Trade-off
- 3 Norm grows with # agents

One interpretation: penalize constraint violation during transients

Augmented Lagrangian  

$$L_{\rho}(x,\nu) = \frac{1}{2}x^{\mathsf{T}}Qx + c^{\mathsf{T}}x + \nu^{\mathsf{T}}(Sx - b) + \frac{\rho}{2}||Sx - b||_{2}^{2}$$

Augmented Primal-Dual System  

$$au_x \dot{x} = -Qx - S^T \nu - c - \rho S^T (Sx - b)$$
  
 $au_\nu \dot{\nu} = Sx - b$ 

One interpretation: penalize constraint violation during transients

Augmented Lagrangian  

$$L_{\rho}(x,\nu) = \frac{1}{2}x^{\mathsf{T}}Qx + c^{\mathsf{T}}x + \nu^{\mathsf{T}}(Sx - b) + \frac{\rho}{2}||Sx - b||_{2}^{2}$$

Augmented Primal-Dual System  $\tau_{x}\dot{x} = -Qx - S^{\mathsf{T}}\nu - c - \rho S^{\mathsf{T}}(Sx - b)$   $\tau_{\nu}\dot{\nu} = Sx - b$ 

One interpretation: penalize constraint violation during transients

Augmented Lagrangian  

$$L_{\rho}(x,\nu) = \frac{1}{2}x^{\mathsf{T}}Qx + c^{\mathsf{T}}x + \nu^{\mathsf{T}}(Sx - b) + \frac{\rho}{2}||Sx - b||_{2}^{2}$$

Augmented Primal-Dual System  

$$G_{\rho}: \begin{cases} \tau_{x} \dot{x} = -Qx - c - S^{\mathsf{T}} \nu - \rho S^{\mathsf{T}} (Sx - b) + d_{c} \\ \tau_{\nu} \dot{\nu} = Sx - b + d_{b} \\ z = \frac{1}{\sqrt{2}} Qx \end{cases}$$

One interpretation: penalize constraint violation during transients

Augmented Lagrangian  

$$L_{\rho}(x,\nu) = \frac{1}{2}x^{\mathsf{T}}Qx + c^{\mathsf{T}}x + \nu^{\mathsf{T}}(Sx - b) + \frac{\rho}{2}||Sx - b||_{2}^{2}$$

Augmented Primal-Dual System  

$$G_{\rho}:\begin{cases} \tau_{x}\dot{x} = -Qx - c - S^{\mathsf{T}}\nu - \rho S^{\mathsf{T}}(Sx - b) + d_{c} \\ \tau_{\nu}\dot{\nu} = Sx - b + d_{b} \\ z = \frac{1}{\sqrt{2}}Qx \end{cases}$$

Uniformity Assumption:  $Q_{ii} = q$ ,  $\tau_{x,ii} = \tau_x$ ,  $\tau_{\nu,ii} = \tau_{\nu}$ .

## Block Diagram of Primal-Dual I/O System



**Theorem**: Aug. Primal-Dual Performance The  $\mathcal{H}_2$  norm of the augmented P-D system is

$$\|G_{\rho}\|_{\mathcal{H}_{2}}^{2} = \frac{1}{4\tau_{x}}(n-r) + \frac{1}{4}\left(\frac{1}{\tau_{x}} + \frac{1}{\tau_{\nu}}\right)\sum_{k=1}^{r}\frac{q}{q+\rho\,\sigma_{k}(S)}$$

where  $\sigma_k(S) = \text{singular values}$ . Moreover

$$\frac{\|\mathsf{\textit{G}}_{\rho}\|_{\mathcal{H}_{2}} \leq \|\mathsf{\textit{G}}\|_{\mathcal{H}_{2}}}{\text{power and }} \quad \text{and} \quad \left|\lim_{\rho \to \infty} \|\mathsf{\textit{G}}_{\rho}\|_{\mathcal{H}_{2}}^{2} = \frac{1}{4\tau_{x}}(n-r)\right|$$



## **Key Insights**

- **()** Not independent of Q, S
- More constraints = good
- More gain = good

**Theorem**: Aug. Primal-Dual Performance The  $\mathcal{H}_2$  norm of the augmented P-D system is

$$\|G_{\rho}\|_{\mathcal{H}_{2}}^{2} = \frac{1}{4\tau_{x}}(n-r) + \frac{1}{4}\left(\frac{1}{\tau_{x}} + \frac{1}{\tau_{\nu}}\right)\sum_{k=1}^{r}\frac{q}{q+\rho\,\sigma_{k}(S)}$$

where  $\sigma_k(S) = \text{singular values}$ . Moreover

$$\boxed{\|G_\rho\|_{\mathcal{H}_2} \leq \|G\|_{\mathcal{H}_2}} \quad \text{ and } \quad \left|\lim_{\rho \to \infty} \|G_\rho\|_{\mathcal{H}_2}^2 = \frac{1}{4\tau_x}(n-r)\right|$$

Recall:
$$\|G\|_{\mathcal{H}_2}^2 = \frac{1}{4\tau_x}n + \frac{1}{4\tau_\nu}r$$

## **Key Insights**

- **O** Not independent of Q, S
- More constraints = good
- O More gain = good

**Theorem**: Aug. Primal-Dual Performance The  $\mathcal{H}_2$  norm of the augmented P-D system is

$$\|G_{\rho}\|_{\mathcal{H}_{2}}^{2} = \frac{1}{4\tau_{x}}(n-r) + \frac{1}{4}\left(\frac{1}{\tau_{x}} + \frac{1}{\tau_{\nu}}\right)\sum_{k=1}^{r}\frac{q}{q+\rho\,\sigma_{k}(S)}$$

where  $\sigma_k(S) = \text{singular values}$ . Moreover

$$\boxed{\| {\it G}_\rho \|_{{\cal H}_2} \leq \| {\it G} \|_{{\cal H}_2}} \qquad {\rm and} \qquad \left| {\mathop {\lim }\limits_{\rho \to \infty } \| {\it G}_\rho \|_{{\cal H}_2}^2 = \frac{1}{4\tau_{\scriptscriptstyle X}}(n-r)} \right|$$

Recall:
$$\|G\|_{\mathcal{H}_2}^2 = \frac{1}{4\tau_x}n + \frac{1}{4\tau_\nu}r$$

#### **Key Insights**

- **O** Not independent of Q, S
- O More constraints = good

More gain = good

**Theorem**: Aug. Primal-Dual Performance The  $\mathcal{H}_2$  norm of the augmented P-D system is

$$\|G_{\rho}\|_{\mathcal{H}_{2}}^{2} = \frac{1}{4\tau_{x}}(n-r) + \frac{1}{4}\left(\frac{1}{\tau_{x}} + \frac{1}{\tau_{\nu}}\right)\sum_{k=1}^{r}\frac{q}{q+\rho\,\sigma_{k}(S)}$$

where  $\sigma_k(S) = \text{singular values}$ . Moreover

$$\frac{\|G_{\rho}\|_{\mathcal{H}_2} \leq \|G\|_{\mathcal{H}_2}}{\text{and}} \qquad \text{and} \qquad \left|\lim_{\rho \to \infty} \|G_{\rho}\|_{\mathcal{H}_2}^2 = \frac{1}{4\tau_x}(n-r)\right|$$

Recall:  
$$\|G\|_{\mathcal{H}_2}^2 = \frac{1}{4\tau_x}n + \frac{1}{4\tau_\nu}r$$

## **Key Insights**

- **O** Not independent of Q, S
- O More constraints = good
- More gain = good

# Optimal Frequency Regulation Problem minimize $\sum_{i=1}^{n} \frac{1}{2}k_i p_i^2$ subject to $\sum_{i=1}^{n} (P_i^* + p_i) = 0$

After dual decomposition:  $P_i^* = c_i =$  uncontrolled load

 $\implies$  Quantify  $\mathcal{H}_2$ -norm from **load disturbance**  $d_c$  to **generation cost** 

1 For Primal-Dual : 
$$\|G\|_{\mathcal{H}_2} \propto \sqrt{n}$$

 $\Rightarrow$  Large systems are more sensitive

- 3 For Augmented Primal-Dual:  $\|G_{\rho}\|_{\mathcal{H}_2}^2 \to \text{const.} \text{ as } \rho \to \infty$ 
  - $\implies$  Sensitivity removed by high-gain (Laplacian!!) feedback

#### Optimal Frequency Regulation Problem

$$\begin{array}{l} \underset{p \in \mathbb{R}^n}{\text{minimize}} \quad \sum_{i=1}^n \frac{1}{2} k_i p_i^2 \\ \text{subject to} \sum_{i=1}^n (P_i^* + p_i) = 0 \end{array}$$

After dual decomposition:  $P_i^* = c_i =$ uncontrolled load

 $\implies$  Quantify  $\mathcal{H}_2$ -norm from load disturbance  $d_c$  to generation cost

• For Primal-Dual :  $\|G\|_{\mathcal{H}_2} \propto \sqrt{n}$ 

 $\Rightarrow$  Large systems are more sensitive

- 3 For Augmented Primal-Dual:  $\|G_{\rho}\|_{\mathcal{H}_2}^2 \to \text{const.} \text{ as } \rho \to \infty$ 
  - $\implies$  Sensitivity removed by high-gain (Laplacian!!) feedback

#### Optimal Frequency Regulation Problem

$$\begin{array}{l} \underset{p \in \mathbb{R}^n}{\text{minimize}} \quad \sum_{i=1}^n \frac{1}{2} k_i p_i^2 \\ \text{subject to} \sum_{i=1}^n (P_i^* + p_i) = 0 \end{array}$$

After dual decomposition:  $P_i^* = c_i =$ uncontrolled load

 $\implies$  Quantify  $\mathcal{H}_2$ -norm from load disturbance  $d_c$  to generation cost

• For Primal-Dual :  $\|G\|_{\mathcal{H}_2} \propto \sqrt{n}$ 

⇒ Large systems are more sensitive

3 For Augmented Primal-Dual:  $\|G_{\rho}\|_{\mathcal{H}_2}^2 \to \text{const.} \text{ as } \rho \to \infty$ 

 $\implies$  Sensitivity removed by high-gain (Laplacian!!) feedback

#### Optimal Frequency Regulation Problem

$$\begin{array}{l} \underset{p \in \mathbb{R}^n}{\text{minimize}} \quad \sum_{i=1}^n \frac{1}{2} k_i p_i^2 \\ \text{subject to} \sum_{i=1}^n (P_i^* + p_i) = 0 \end{array}$$

After dual decomposition:  $P_i^* = c_i =$ uncontrolled load

 $\implies$  Quantify  $\mathcal{H}_2$ -norm from load disturbance  $d_c$  to generation cost

• For Primal-Dual : 
$$\|G\|_{\mathcal{H}_2} \propto \sqrt{n}$$

 $\implies$  Large systems are more sensitive

- 3 For Augmented Primal-Dual:  $\|G_{\rho}\|_{\mathcal{H}_2}^2 \to \text{const.}$  as  $\rho \to \infty$ 
  - $\implies$  Sensitivity removed by high-gain (Laplacian!!) feedback

#### Optimal Frequency Regulation Problem

$$\begin{array}{l} \underset{p \in \mathbb{R}^n}{\text{minimize}} \quad \sum_{i=1}^n \frac{1}{2} k_i p_i^2 \\ \text{subject to} \sum_{i=1}^n (P_i^* + p_i) = 0 \end{array}$$

After dual decomposition:  $P_i^* = c_i =$ uncontrolled load

 $\implies$  Quantify  $\mathcal{H}_2$ -norm from load disturbance  $d_c$  to generation cost

• For Primal-Dual : 
$$||G||_{\mathcal{H}_2} \propto \sqrt{n}$$

 $\implies$  Large systems are more sensitive

**2** For Augmented Primal-Dual:  $\|G_{\rho}\|_{\mathcal{H}_{2}}^{2} \rightarrow \text{const. as } \rho \rightarrow \infty$ 

 $\Rightarrow$  Sensitivity removed by high-gain (Laplacian!!) feedback

#### Optimal Frequency Regulation Problem

$$\begin{array}{l} \underset{p \in \mathbb{R}^n}{\text{minimize}} \quad \sum_{i=1}^n \frac{1}{2} k_i p_i^2 \\ \text{subject to} \sum_{i=1}^n (P_i^* + p_i) = 0 \end{array}$$

After dual decomposition:  $P_i^* = c_i =$ uncontrolled load

 $\implies$  Quantify  $\mathcal{H}_2$ -norm from load disturbance  $d_c$  to generation cost

• For Primal-Dual : 
$$\|G\|_{\mathcal{H}_2} \propto \sqrt{n}$$

 $\implies$  Large systems are more sensitive

- **2** For Augmented Primal-Dual:  $\|G_{\rho}\|_{\mathcal{H}_2}^2 \to \text{const.}$  as  $\rho \to \infty$ 
  - $\implies$  Sensitivity removed by high-gain (Laplacian!!) feedback

# Conclusions

Primal-Dual Input/Output Performance:

- **1**  $\mathcal{H}_2$ -norm scales with  $\sqrt{\#}$  agents
- ② Augmentation reduces  $\mathcal{H}_2$  norm, constraints are good
- Implications for frequency control of power systems

For  $L_2$ -gain or  $\mathcal{H}_{\infty}$  results, see Allerton 2016 paper "Input/Output Analysis of Primal-Dual Gradient Algorithms"

#### What's next?

- Inequality constraints? Descriptor systems?
- Other distributed optimization algorithms
- Beyond augmented Lagrangians ...

# Conclusions

Primal-Dual Input/Output Performance:

- **1**  $\mathcal{H}_2$ -norm scales with  $\sqrt{\#}$  agents
- ② Augmentation reduces  $\mathcal{H}_2$  norm, constraints are good
- Implications for frequency control of power systems

For  $L_2$ -gain or  $\mathcal{H}_{\infty}$  results, see Allerton 2016 paper "Input/Output Analysis of Primal-Dual Gradient Algorithms"

#### What's next?

- Inequality constraints? Descriptor systems?
- Other distributed optimization algorithms
- Beyond augmented Lagrangians . . .

## Acknowledgements



Bala K. Poolla



Nima Monshizadeh



Florian Dörfler



Eidgenössische Technische Hochschule Zürich Swiss Federal Institute of Technology Zurich





## Question Time



https://ece.uwaterloo.ca/~jwsimpso/ jwsimpson@uwaterloo.ca

**1** Simplify: Assume block-diagonal solution to avoid coupling

O Lyapunov Equation:

$$\begin{bmatrix} Y_{11} & 0 \\ 0 & Y_{22} \end{bmatrix} \begin{bmatrix} -\tau_x^{-1}Q & -\tau_x^{-1}S^{\mathsf{T}} \\ \tau_\nu^{-1}S & 0 \end{bmatrix} + \begin{bmatrix} \star & \star \\ \star & \star \end{bmatrix}^{\mathsf{T}} = -\begin{bmatrix} \frac{1}{2}Q & 0 \\ 0 & 0 \end{bmatrix}$$

$$Y_{11}\tau_x^{-1}Q + Q\tau_x^{-1}Y_{11} - \frac{1}{2}Q = 0, \implies Y_{11} = \tau_x/4$$
  
$$Y_{22}\tau_\nu^{-1}S - S\tau_x^{-1}Y_{11} = 0, \implies Y_{22} = S\tau_x^{-1}Y_{11}S^{\dagger}\tau_\nu$$
  
$$= SS^{\dagger}\tau_\nu/4$$
  
$$= \tau_\nu/4$$

**3** Trace Formula:  $\|G\|_{\mathcal{H}_2}^2 = \text{Tr}(\tau_x^{-1}Y_{11}\tau_x^{-1}) + (\star)$ 

- **1** Simplify: Assume block-diagonal solution to avoid coupling
- **O** Lyapunov Equation:

$$\begin{bmatrix} Y_{11} & 0 \\ 0 & Y_{22} \end{bmatrix} \begin{bmatrix} -\tau_x^{-1}Q & -\tau_x^{-1}S^\mathsf{T} \\ \tau_\nu^{-1}S & 0 \end{bmatrix} + \begin{bmatrix} \star & \star \\ \star & \star \end{bmatrix}^\mathsf{T} = -\begin{bmatrix} \frac{1}{2}Q & 0 \\ 0 & 0 \end{bmatrix}$$

$$Y_{11}\tau_{x}^{-1}Q + Q\tau_{x}^{-1}Y_{11} - \frac{1}{2}Q = 0, \implies Y_{11} = \tau_{x}/4$$
$$Y_{22}\tau_{\nu}^{-1}S - S\tau_{x}^{-1}Y_{11} = 0, \implies Y_{22} = S\tau_{x}^{-1}Y_{11}S^{\dagger}\tau_{\nu}$$
$$= SS^{\dagger}\tau_{\nu}/4$$
$$= \tau_{\nu}/4$$

**3** Trace Formula:  $\|G\|_{\mathcal{H}_2}^2 = \text{Tr}(\tau_x^{-1}Y_{11}\tau_x^{-1}) + (\star)$ 

- **1** Simplify: Assume block-diagonal solution to avoid coupling
- **O** Lyapunov Equation:

$$\begin{bmatrix} Y_{11} & 0\\ 0 & Y_{22} \end{bmatrix} \begin{bmatrix} -\tau_x^{-1}Q & -\tau_x^{-1}S^{\mathsf{T}}\\ \tau_\nu^{-1}S & 0 \end{bmatrix} + \begin{bmatrix} \star & \star\\ \star & \star \end{bmatrix}^{\mathsf{T}} = -\begin{bmatrix} \frac{1}{2}Q & 0\\ 0 & 0 \end{bmatrix}$$
$$Y_{11}\tau_x^{-1}Q + Q\tau_x^{-1}Y_{11} - \frac{1}{2}Q = 0, \implies Y_{11} = \tau_x/4$$
$$Y_{22}\tau_\nu^{-1}S - S\tau_x^{-1}Y_{11} = 0, \implies Y_{22} = S\tau_x^{-1}Y_{11}S^{\dagger}\tau_1$$
$$= SS^{\dagger}\tau_\nu/4$$

**3** Trace Formula:  $||G||_{\mathcal{H}_2}^2 = \text{Tr}(\tau_x^{-1}Y_{11}\tau_x^{-1}) + (\star)$ 

- **1** Simplify: Assume block-diagonal solution to avoid coupling
- **O** Lyapunov Equation:

$$\begin{bmatrix} Y_{11} & \mathbb{O} \\ \mathbb{O} & Y_{22} \end{bmatrix} \begin{bmatrix} -\tau_x^{-1}Q & -\tau_x^{-1}S^{\mathsf{T}} \\ \tau_\nu^{-1}S & \mathbb{O} \end{bmatrix} + \begin{bmatrix} \star & \star \\ \star & \star \end{bmatrix}^{\mathsf{T}} = -\begin{bmatrix} \frac{1}{2}Q & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{bmatrix}$$

$$Y_{11}\tau_{x}^{-1}Q + Q\tau_{x}^{-1}Y_{11} - \frac{1}{2}Q = 0, \implies Y_{11} = \tau_{x}/4$$
$$Y_{22}\tau_{\nu}^{-1}S - S\tau_{x}^{-1}Y_{11} = 0, \implies Y_{22} = S\tau_{x}^{-1}Y_{11}S^{\dagger}\tau_{\nu}$$
$$= SS^{\dagger}\tau_{\nu}/4$$

**3** Trace Formula:  $\|G\|_{\mathcal{H}_2}^2 = \operatorname{Tr}(\tau_x^{-1}Y_{11}\tau_x^{-1}) + (\star)$ 

- **1** Simplify: Assume block-diagonal solution to avoid coupling
- **2** Lyapunov Equation:

$$\begin{bmatrix} Y_{11} & \mathbb{O} \\ \mathbb{O} & Y_{22} \end{bmatrix} \begin{bmatrix} -\tau_x^{-1}Q & -\tau_x^{-1}S^{\mathsf{T}} \\ \tau_\nu^{-1}S & \mathbb{O} \end{bmatrix} + \begin{bmatrix} \star & \star \\ \star & \star \end{bmatrix}^{\mathsf{T}} = -\begin{bmatrix} \frac{1}{2}Q & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{bmatrix}$$

$$Y_{11}\tau_{x}^{-1}Q + Q\tau_{x}^{-1}Y_{11} - \frac{1}{2}Q = 0, \implies Y_{11} = \tau_{x}/4$$
$$Y_{22}\tau_{\nu}^{-1}S - S\tau_{x}^{-1}Y_{11} = 0, \implies Y_{22} = S\tau_{x}^{-1}Y_{11}S^{\dagger}\tau_{\nu}$$
$$= SS^{\dagger}\tau_{\nu}/4$$
$$= \tau_{\nu}/4$$

**3** Trace Formula:  $||G||_{\mathcal{H}_2}^2 = \text{Tr}(\tau_x^{-1}Y_{11}\tau_x^{-1}) + (\star)$ 

- **1** Simplify: Assume block-diagonal solution to avoid coupling
- **2** Lyapunov Equation:

$$\begin{bmatrix} Y_{11} & \mathbb{O} \\ \mathbb{O} & Y_{22} \end{bmatrix} \begin{bmatrix} -\tau_x^{-1}Q & -\tau_x^{-1}S^{\mathsf{T}} \\ \tau_\nu^{-1}S & \mathbb{O} \end{bmatrix} + \begin{bmatrix} \star & \star \\ \star & \star \end{bmatrix}^{\mathsf{T}} = -\begin{bmatrix} \frac{1}{2}Q & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{bmatrix}$$

$$\begin{aligned} Y_{11}\tau_x^{-1}Q + Q\tau_x^{-1}Y_{11} - \frac{1}{2}Q &= 0, \implies Y_{11} &= \tau_x/4 \\ Y_{22}\tau_\nu^{-1}S - S\tau_x^{-1}Y_{11} &= 0, \implies Y_{22} &= S\tau_x^{-1}Y_{11}S^{\dagger}\tau_\nu \\ &= SS^{\dagger}\tau_\nu/4 \\ &= \tau_\nu/4 \end{aligned}$$

**3** Trace Formula:  $\|G\|_{\mathcal{H}_2}^2 = \text{Tr}(\tau_x^{-1}Y_{11}\tau_x^{-1}) + (\star)$ 

- **1** Simplify: Assume block-diagonal solution to avoid coupling
- **O** Lyapunov Equation:

$$\begin{bmatrix} Y_{11} & \mathbb{O} \\ \mathbb{O} & Y_{22} \end{bmatrix} \begin{bmatrix} -\tau_x^{-1}Q & -\tau_x^{-1}S^{\mathsf{T}} \\ \tau_\nu^{-1}S & \mathbb{O} \end{bmatrix} + \begin{bmatrix} \star & \star \\ \star & \star \end{bmatrix}^{\mathsf{T}} = -\begin{bmatrix} \frac{1}{2}Q & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{bmatrix}$$

$$\begin{aligned} Y_{11}\tau_x^{-1}Q + Q\tau_x^{-1}Y_{11} - \frac{1}{2}Q &= 0, \implies Y_{11} &= \tau_x/4 \\ Y_{22}\tau_\nu^{-1}S - S\tau_x^{-1}Y_{11} &= 0, \implies Y_{22} &= S\tau_x^{-1}Y_{11}S^{\dagger}\tau_\nu \\ &= SS^{\dagger}\tau_\nu/4 \\ &= \tau_\nu/4 \end{aligned}$$

**3** Trace Formula: 
$$\|G\|_{\mathcal{H}_2}^2 = \text{Tr}(\tau_x^{-1}Y_{11}\tau_x^{-1}) + (\star)$$

Smart Grid Project Samples Distributed Inverter Control



#### Voltage Collapse (Nat. Comms.)



Optimal Distrib. Volt/Var (CDC)



#### Wide-Area Monitoring (TSG)



appendix

## An incomplete literature review of a busy field

#### ntwk with unknown disturbances $\cup$ integral control $\cup$ distributed averaging

- all-to-all source frequency & injection averaging [Q. Shafiee, J. Vasquez, & J. Guerrero, '13] & [H. Liang, B. Choi, W. Zhuang, & X. Shen, '13] & [M. Andreasson, D. V. Dimarogonas, K. H. Johansson, & H. Sandberg, '12]
- optimality w.r.t. economic dispatch [E. Mallada & S. Low, '13] & [M. Andreasson, D. V. Dimarogonas, K. H. Johansson, & H. Sandberg, '13] & [X. Zhang and A. Papachristodoulou, '13] & [N. Li, L. Chen, C. Zhao & S. Low '13]
- ratio consensus & dispatch [S.T. Cady, A. Garcia-Dominguez, & C.N. Hadjicostis, '13]
- load balancing in Port-Hamiltonian networks [J. Wei & A. Van der Schaft, '13]
- passivity-based network cooperation and flow optimization [M. Bürger, D. Zelazo, & F. Allgöwer, '13, M. Bürger & C. de Persis '13, He Bai & S.Y. Shafi '13]
- distributed PI avg optimization [G. Droge, H. Kawashima, & M. Egerstedt, '13]
- PI avg consensus [R. Freeman, P. Yang, & K. Lynch '06] & [M. Zhu & S. Martinez '10]
- decentralized "practical" integral control [N. Ainsworth & S. Grijalva, '13]

 $egin{aligned} &\omega_i = \omega^* - m_i P_i( heta) - \Omega_i \ &k_i \dot{\Omega}_i = (\omega_i - \omega^*) - \sum_{j \ \subseteq \ \mathrm{inverters}} a_{ij} \cdot (\Omega_i - \Omega_j) \end{aligned}$ 

no tuning, no model dependence

- eak comm. requirements
- 3 maintains load sharing (share burden of sec. control)

Simple & Intuitive

Theorem: Stability of DAPI [JWSP, FD, & FB, '13] DAPI-Controlled System Stable \$\$ Droop-Controlled System Stable

(grid-conscious sec. control)

$$\omega_i = \omega^* - m_i P_i(\theta) - \Omega_i$$
  
$$k_i \dot{\Omega}_i = (\omega_i - \omega^*) - \sum_{j \subseteq \text{inverters}} a_{ij} \cdot (\Omega_i - \Omega_j)$$



no tuning, no model dependence

- eak comm. requirements
- 3 maintains load sharing (share burden of sec. control)

#### Simple & Intuitive

Theorem: Stability of DAPI [JWSP, FD, & FB, '13] DAPI-Controlled System Stable Droop-Controlled System Stable

grid-conscious sec. control)

$$\omega_i = \omega^* - m_i P_i(\theta) - \Omega_i$$
  
$$k_i \dot{\Omega}_i = (\omega_i - \omega^*) - \sum_{j \subseteq \text{inverters}} a_{ij} \cdot (\Omega_i - \Omega_j)$$



- no tuning, no model dependence
- weak comm. requirements
- 6 maintains load sharing

(share burden of sec. control)

#### Simple & Intuitive

Theorem: Stability of DAPI [JWSP, FD, & FB, '13] DAPI-Controlled System Stable Droop-Controlled System Stable

$$\omega_i = \omega^* - m_i P_i(\theta) - \Omega_i$$
  
 $k_i \dot{\Omega}_i = (\omega_i - \omega^*) - \sum_{j \subseteq \text{inverters}} a_{ij} \cdot (\Omega_i - \Omega_j)$ 



no tuning, no model dependence

- weak comm. requirements
- 8 maintains load sharing

(share burden of sec. control)

#### Simple & Intuitive

Theorem: Stability of DAPI [JWSP, FD, & FB, '13] DAPI-Controlled System Stable \$\$\$ Droop-Controlled System Stable

(grid-conscious sec. control)

$$\omega_i = \omega^* - m_i P_i(\theta) - \Omega_i$$
  
 $k_i \dot{\Omega}_i = (\omega_i - \omega^*) - \sum_{j \subseteq \text{inverters}} a_{ij} \cdot (\Omega_i - \Omega_j)$ 



no tuning, no model dependence

- weak comm. requirements
- 8 maintains load sharing

(share burden of sec. control)

#### Simple & Intuitive

Theorem: Stability of DAPI [JWSP, FD, & FB, '13] DAPI-Controlled System Stable \$\$\$ Droop-Controlled System Stable

(grid-conscious sec. control)

## From Hierarchical Control to Distributed Control

flat hierarchy, no time-scale separations, & model-free



# From Hierarchical Control to Distributed Control

flat hierarchy, no time-scale separations, & model-free

