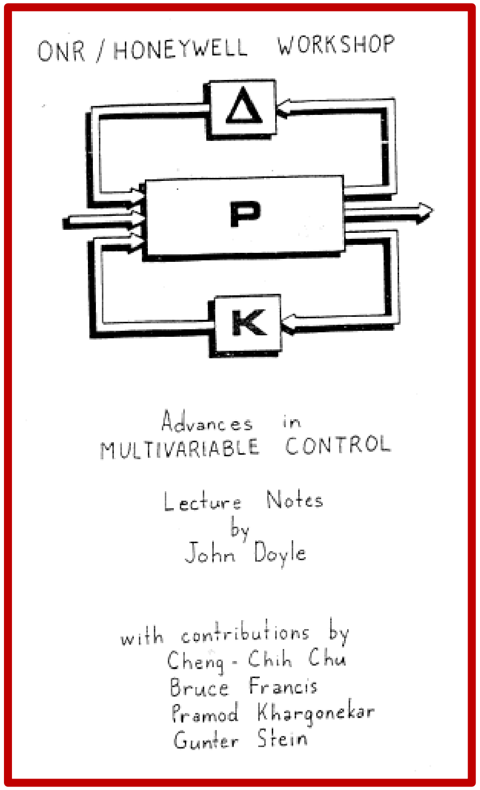
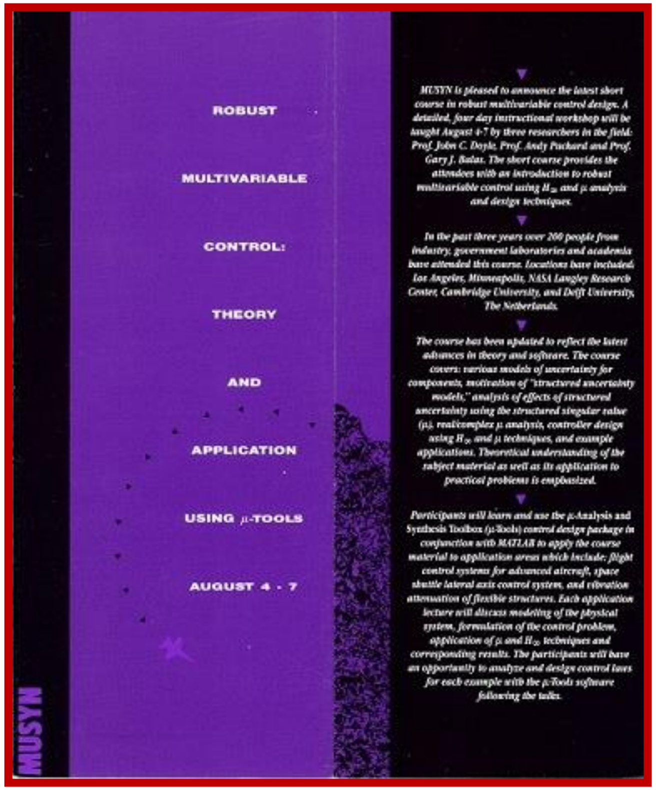
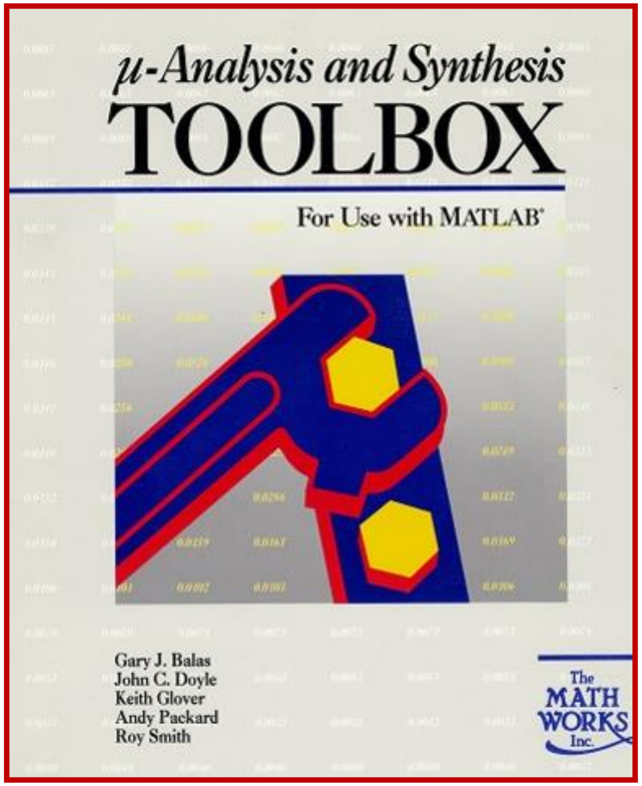
# Robust Control: Past Successes and Future Directions


**Peter Seiler**
**University of Minnesota**


**February 6, 2019**



# Outline

- **Brief Overview of Robust Control**
- Robustness of Time-Varying Systems
- Future Directions
- Conclusions

# Pillars of Robust Control

1. Multivariable Optimal Control
   - $H_2$, $H_\infty$, DK-synthesis

2. Fundamental Limitations of Dynamics & Control
   - Bode sensitivity integral, complementary sensitivity integrals, constraints due to right-half plane poles and zeros.

3. Uncertainty Modeling and Robustness Analysis
   - Linear Fractional Transformations (LFTs), Structured Singular Value ($\mu$), Integral Quadratic Constraints (IQCs)

# Pillars of Robust Control

1. Multivariable Optimal Control
   - $H_2$, $H_\infty$, DK-synthesis

2. **Fundamental Limitations of Dynamics & Control**
   - Bode sensitivity integral, complementary sensitivity integrals, constraints due to right-half plane poles and zeros.

3. **Uncertainty Modeling and Robustness Analysis**
   - Linear Fractional Transformations (LFTs), Structured Singular Value ($\mu$), Integral Quadratic Constraints (IQCs)

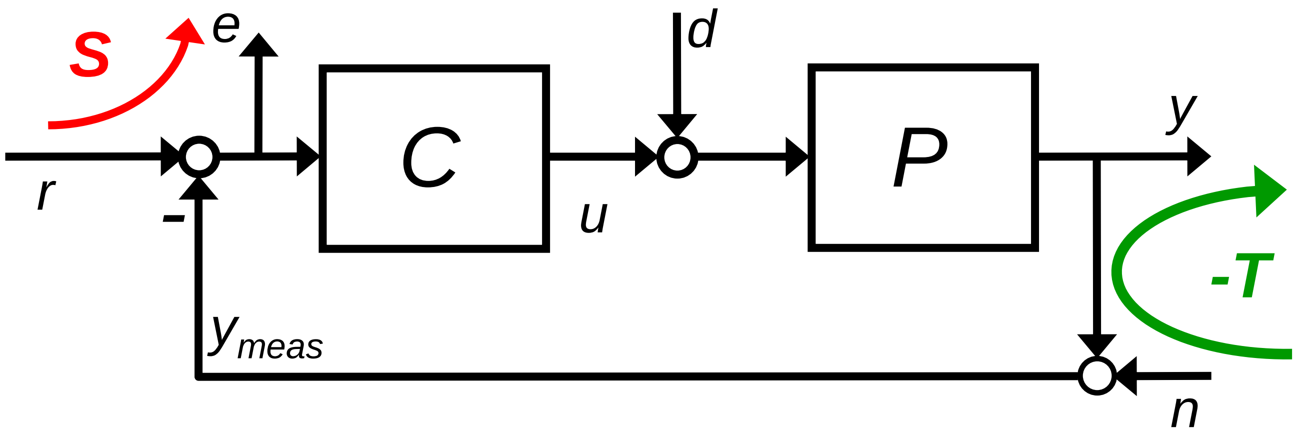
# Basic Feedback Loop

**Sensitivity**

$$S = \frac{1}{1+PC}$$

**Complementary Sensitivity**

$$T = \frac{PC}{1+PC}$$




**Many design objectives:** Stability, disturbance rejection, reference tracking, noise rejection, moderate actuator commands, adequate robustness margins.

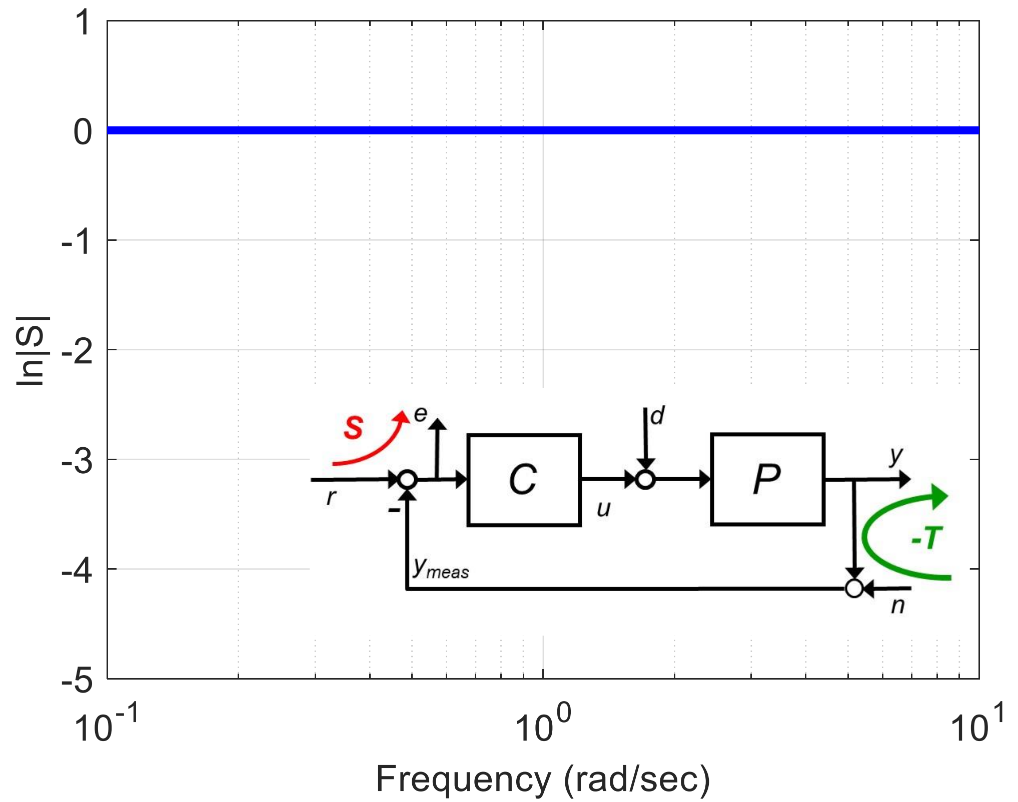
**Basic Limitation:** S+T=1

Typically require |S| ≪ 1 at low frequencies for reference tracking and disturbance rejection.

# Conservation of Sensitivity

Suppose $P$ is stable so that $C = 0$ is a stabilizing controller.

$$S(s) := \frac{1}{1+P(s)C(s)} \equiv 1 \qquad \textbf{AND} \qquad \int_0^{\infty} \ln|S(j\omega)|\, d\omega = 0$$

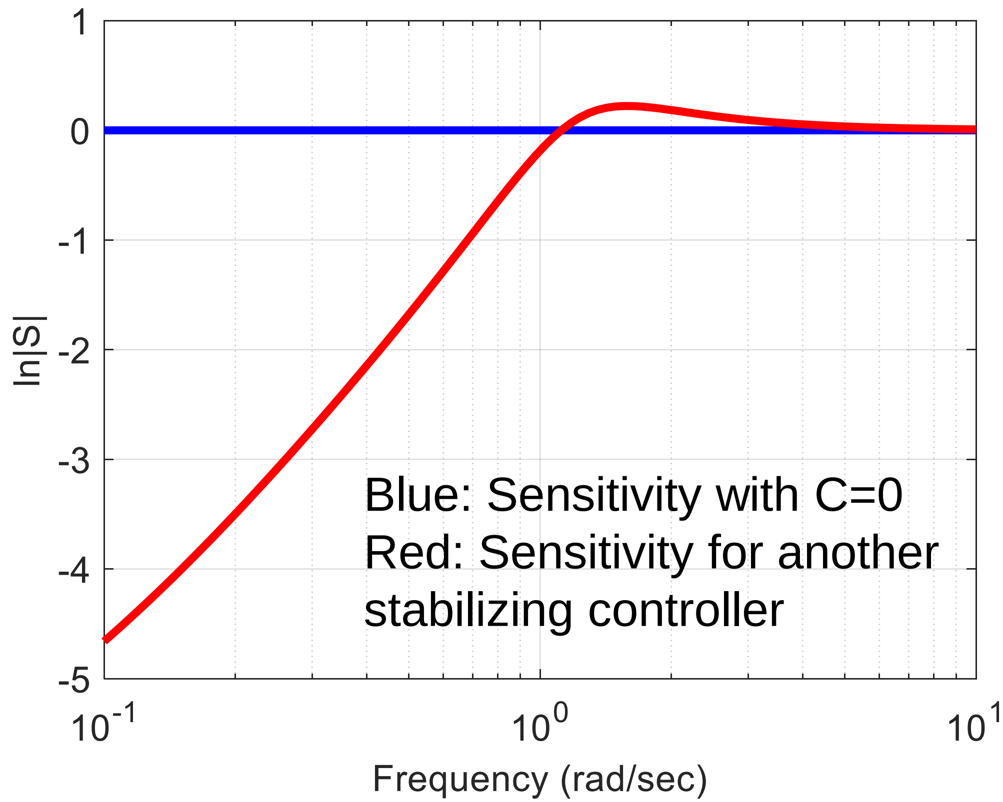


# Conservation of Sensitivity

Suppose *P* is stable so that *C* = 0 is a stabilizing controller.

$$S(s) := \frac{1}{1+P(s)C(s)} \equiv 1 \qquad \textbf{AND} \qquad \int_0^\infty \ln|S(j\omega)|\,d\omega = 0$$

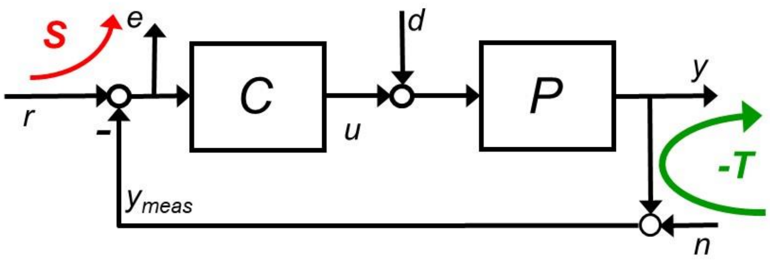
Improving sensitivity at some frequencies leads to degradations at others.



# Bode Integral Theorem [1]

If *PC* is stable, relative degree 2 and *S(s)* is stable. Then:

$$\int_0^{\infty} \ln|S(j\omega)|\,d\omega = 0$$



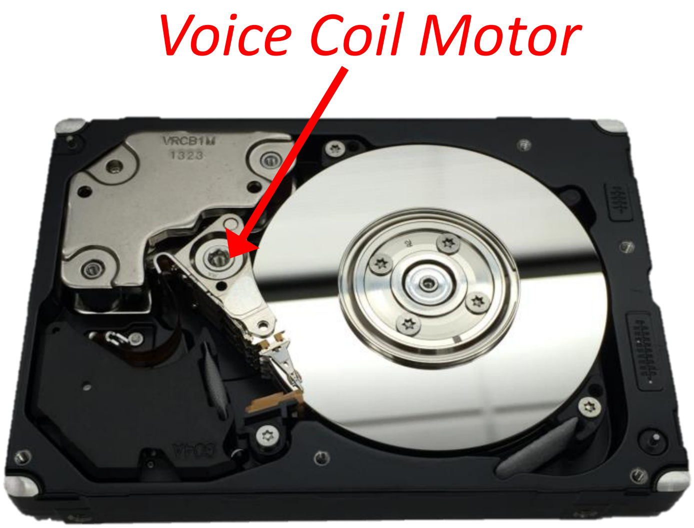

**This a key conserved quantity in feedback design.** Improving performance (e.g. increased bandwidth) comes at the expense of reduced robustness (peak in |*S*|) [2].

Trade-off degrades further if open loop is unstable [3].

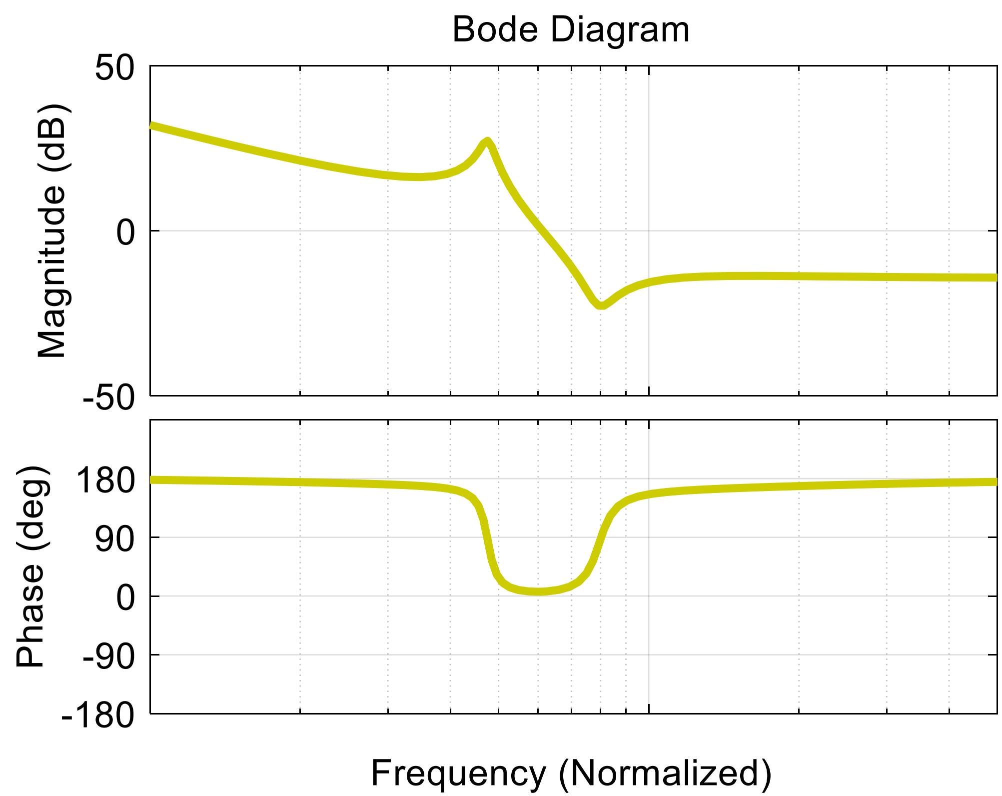

[1] Bode, Network Analysis and Feedback Amplifier Design, 1945.
[2] Stein, Respect the Unstable, Bode Lecture, 1989 (and IEEE CSM, 2003)
[3] Freudenberg and Looze, Frequency Domain Properties of Scalar and Multivariable Feedback Systems, Springer-Verlag, 1988.

# Plant Uncertainty

A simplified model *P* is used for control design.



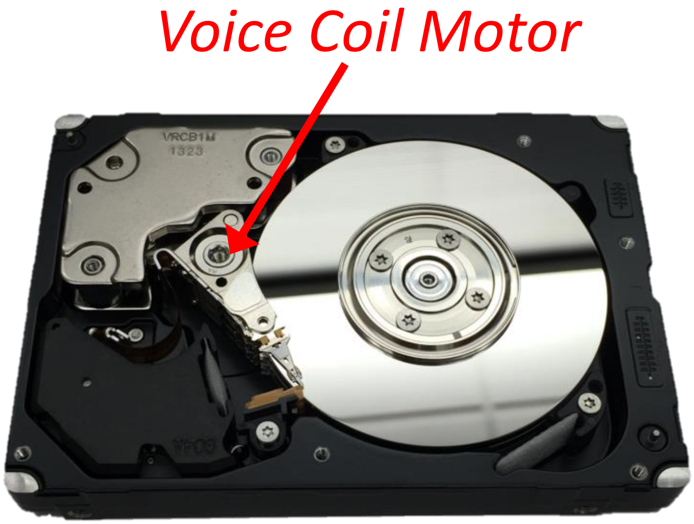

Experimental frequency responses (blue) and simplified model (black).



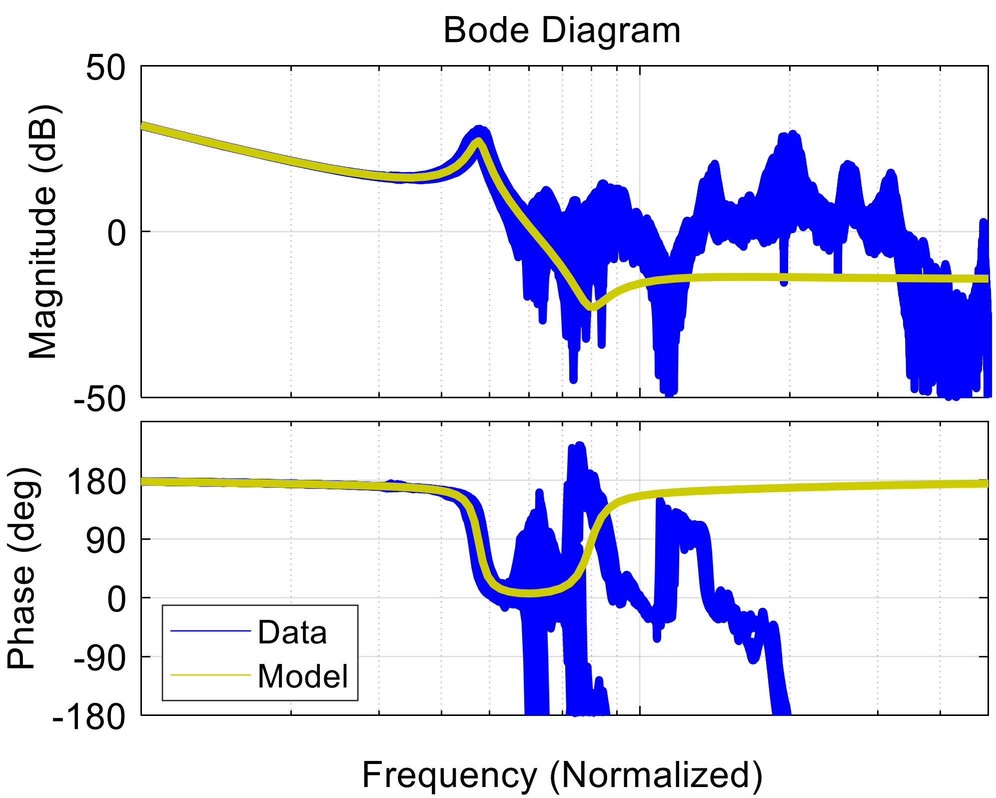
# Plant Uncertainty

A simplified model *P* is used for control design.

- Actual dynamics are complex and have part-to-part variation.
- We lose model fidelity as we go to higher frequencies.




Experimental frequency responses (blue) and simplified model (black).



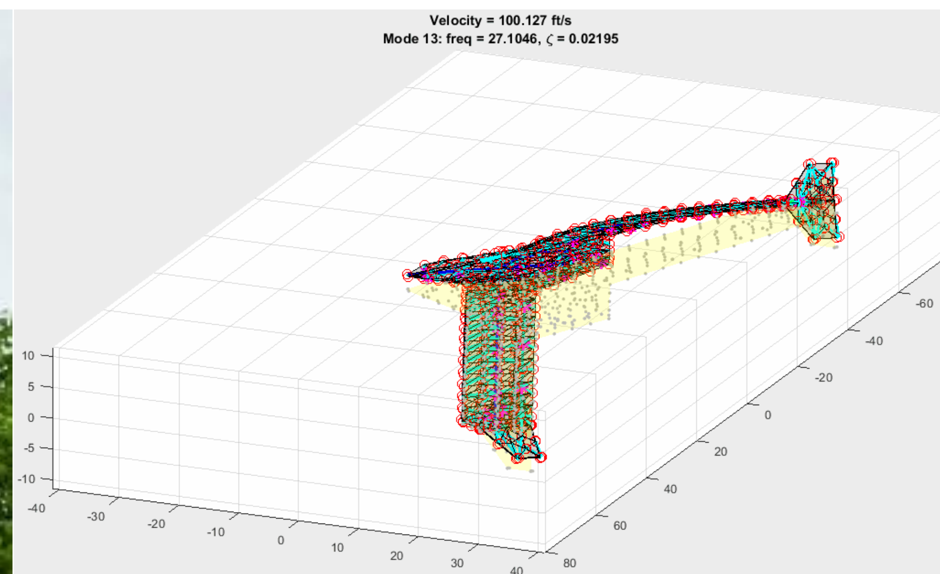
# Stability Margins: Safety Factors for Control

## An Approach:

1. Build an analysis model (possibly of high fidelity)

2. Assess the impact of parametric model errors, e.g. statistical sampling methods or classical gain/phase margins

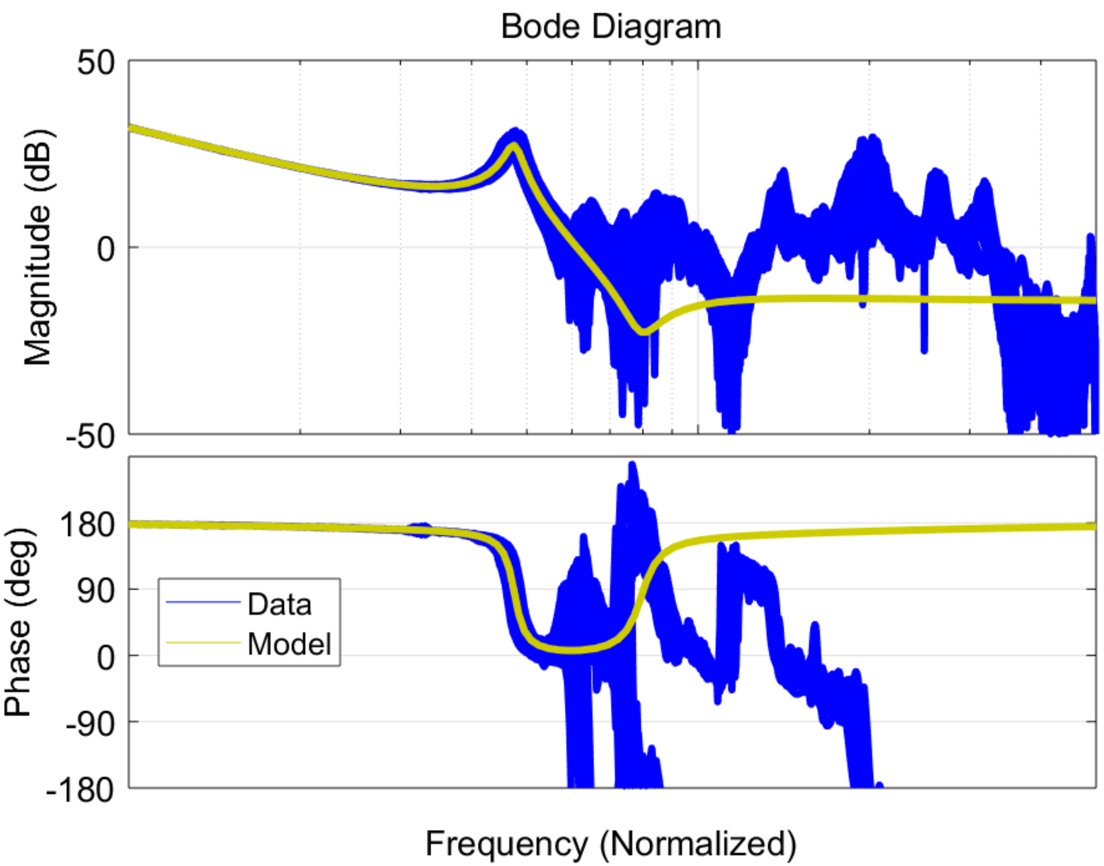
**Issue:** Even high fidelity models fail to capture certain aspects of the dynamics, i.e. there are "unknown unknowns."



# Non-parametric (Dynamic) Uncertainty

Model *nominal* behavior with LTI system $G_0$.

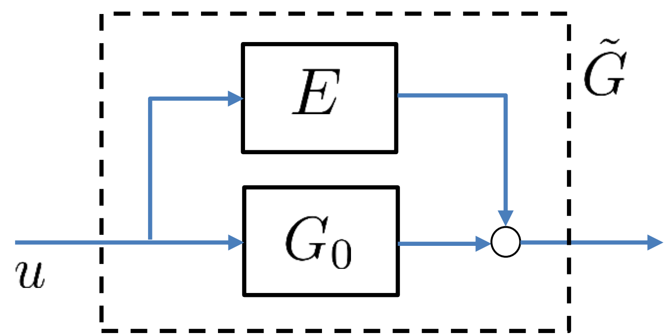
Uncertainty modeled by LTI systems $\tilde{G}$ close to $G_0$ in frequency response, e.g. small additive error.




$$\underbrace{\left|\tilde{G}(\omega) - G_0(\omega)\right|}_{=:\, E} \leq \alpha(\omega)$$

*Error Bound*

$$|E(\omega)| \leq \alpha(\omega)$$



# Advanced Robustness Analysis

Move beyond classical SISO stability (gain/phase) margins

1. Multi-loop (MIMO) systems with multiple uncertainties
2. More detailed uncertainty descriptions including

*Structured Singular Value ($\mu$)* { Parametric, Non-parametric (dynamic) }

- Parametric,
- Non-parametric (dynamic)
- Nonlinearities, e.g. saturation

} *Integral Quadratic Constraints (IQCs)*

3. Consider both robust stability and robust performance

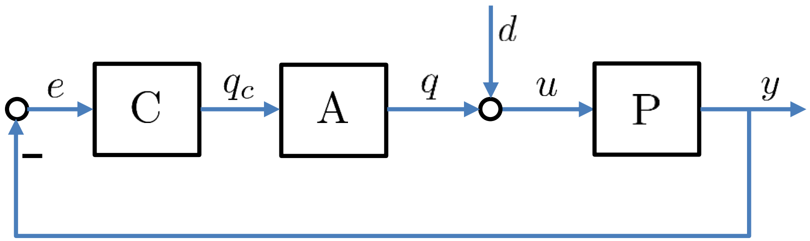
Developments go back to the Lur'e problem (40's) with key contributions in the 80's and 90's:

- μ: Safonov, Stein, Doyle, Packard, …
- IQCs: Yakubovich, Megretski, Rantzer, …

# Numerical Algorithms and Software

Reliable software to create uncertainty models & perform analyses.

- Matlab's Robust Control Toolbox (Safonov & Chiang), (Balas, Doyle, Glover, Packard, & Smith), (Gahinet, Nemirovski, Laub, & Chilali)
- ONERA's Systems Modeling, Analysis and Control Toolbox (Biannic, Burlion, Demourant, Ferreres, Hardier, Loquen, & Roos)




Example Matlab code to assess robustness of simple feedback loop.

```
% Unstable plant with parametric uncertainty
a = ureal('a',1, 'Range', [0.8  1.1]);
b = ureal('b',2, 'Range', [1.7  2.6]);
P = tf(b, [1 -a]);

% Actuator with non-parametric (dynamic) unc.
nomAct = tf(10, [1 10]);
DeltaE = ultidyn('DeltaE',[1 1]);
A = nomAct + 0.1*DeltaE;

% Uncertain closed-loop (d->e) with PI control
C = tf([3 4.5],[1 0]);
R = feedback(-P, A*C);

% Robust stability and worst-case gain
[StabMargin, DestabilizingUncert] = robstab(R);
[wcGain, OffendingUncertainty] = wcgain(R);
```

# Numerical Algorithms and Software

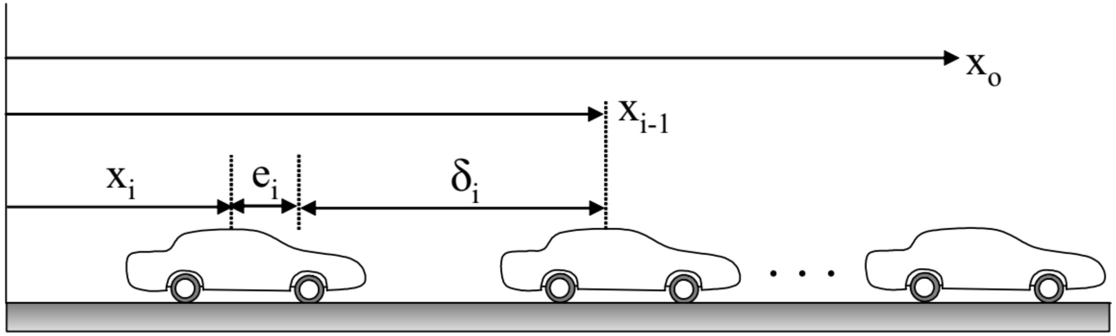
Reliable software to create uncertainty models & perform analyses.

- Matlab's Robust Control Toolbox (Safonov & Chiang), (Balas, Doyle, Glover, Packard, & Smith), (Gahinet, Nemirovski, Laub, & Chilali)
- ONERA's Systems Modeling, Analysis and Control Toolbox (Biannic, Burlion, Demourant, Ferreres, Hardier, Loquen, & Roos)

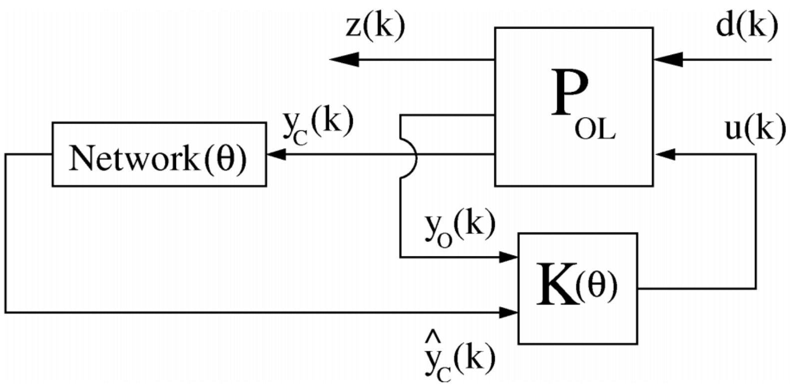
Numerical algorithms continue to be developed, e.g. in Matlab:

- Structured $H_\infty$(R2010b) and `systune` (R2014a): Based on work by (Gahinet, Apkarian, Noll)
- $\mu$ without frequency gridding (R2016b): (Gahinet, Balas, Packard, Seiler) and (Biannic, Ferreres, Roos)
- Automatic regularization for $H_2$ (R2017b) and $H_\infty$ synthesis (R2018b): (Gahinet, Packard, Seiler)
- Multi-loop disk margins (R2018b): (Gahinet, Packard, Seiler)

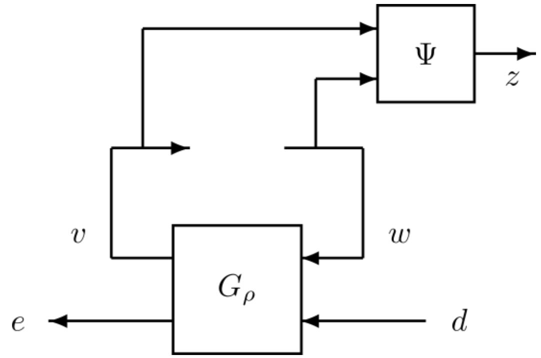
# (My) Theoretical Contributions to Robust Control




Fundamental limits in vehicle platoons
(Seiler, Pant, Hedrick)




Networked Control
(Seiler & Sengupta)

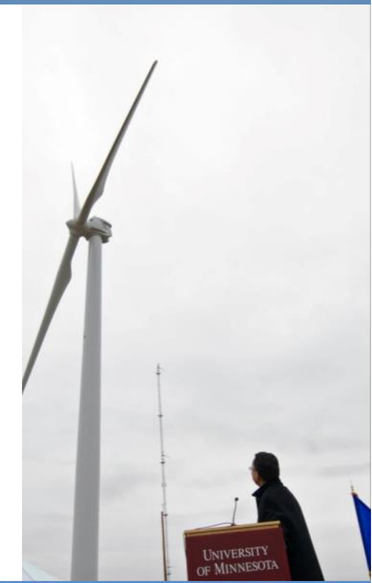


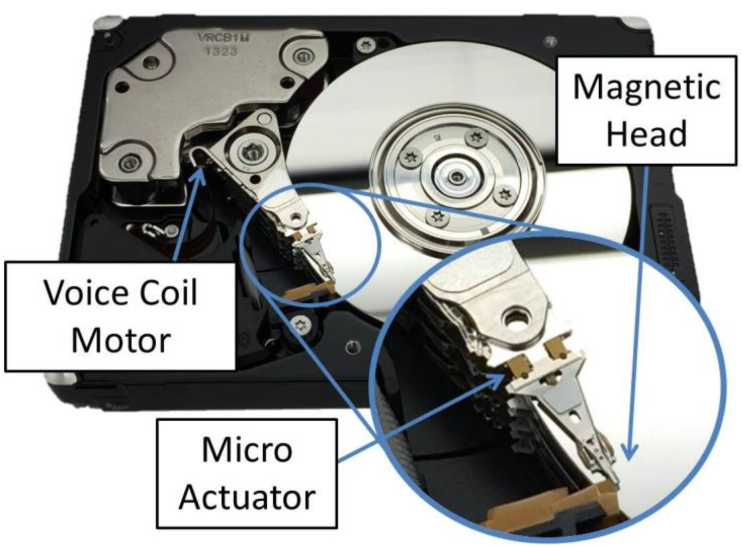

Robustness of Linear Parameter Varying Systems using IQCs
(Seiler, Pfifer, Wang, & Venkataraman)

# (My) Applications of Robust Control


- 787 Flight Control Electronics
- Wind farm modeling and control (Annoni ‘16, Singh, Hoyt)
- Individual turbine control (Wang ‘16, Ossmann, Theis)
- UAV control with a single aerodynamic surface (Venkataraman ’18)
- Flexible aircraft (Kotikalpudi ’17, Theis ’18, Gupta, Pfifer)
- Dual stage hard disk drives with Seagate (Honda ‘16)

(Years refer to Ph.D. theses.)




# Outline

- Brief Overview of Robust Control
- **Robustness of Time-Varying Systems**
  - **Joint work with M. Arcak, A. Packard, M. Moore, and C. Meissen at UC, Berkeley + Jyot Buch at Minnesota.**
  - **Funded by ONR BRC with B. Holm-Hansen at Tech. Monitor**
- Future Directions
- Conclusions

# Time-Varying Systems

**Wind Turbine**
Periodic /
Parameter-Varying

**Flexible Aircraft**
Parameter-Varying


**Vega Launcher**
Time-Varying
(Source: ESA)

**Robotics**
Time-Varying
(Source: ReWalk)

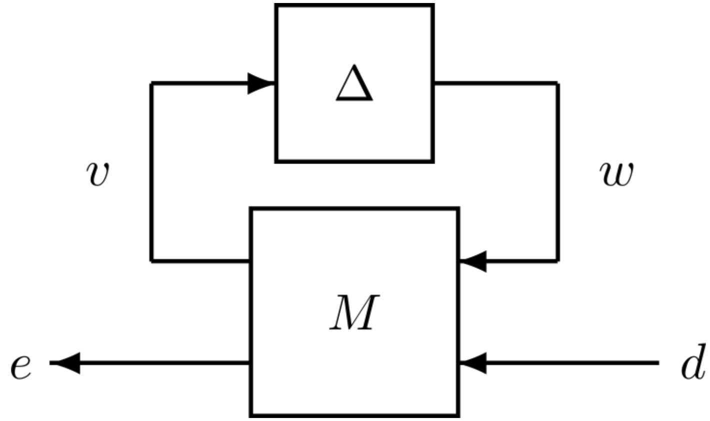
**Few numerically reliable methods to assess the robustness of time-varying systems.**

# (Robust) Finite-Horizon Analysis

## Uncertain LTV System

$$\begin{bmatrix} \dot{x}(t) \\ v(t) \\ e(t) \end{bmatrix} = \begin{bmatrix} A(t) & B_1(t) & B_2(t) \\ C_1(t) & D_1(t) & D_2(t) \\ C_2(t) & 0 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ w(t) \\ d(t) \end{bmatrix}$$

$$x(0) = 0$$



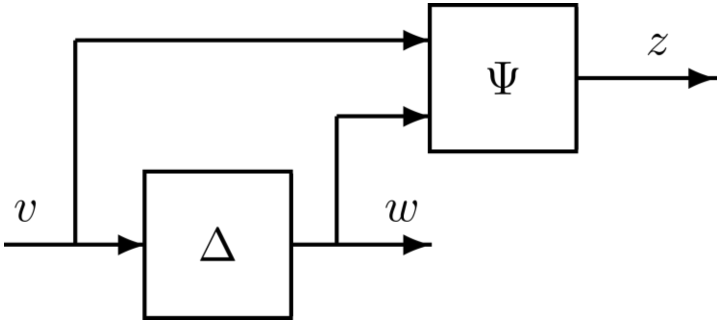

Uncertainty set $\mathbf{\Delta}$ can be block-structured with parametric / non-parametric uncertainties and nonlinearities.

## Analysis Objective

Derive bound on $\|e(T)\|_2$ that holds for all disturbances $\|d\|_{2,[0,T]} \leq 1$ and uncertainties $\Delta \in \mathbf{\Delta}$ on the horizon $[0, T]$.

# Integral Quadratic Constraints (IQCs) [1,2]

The robustness analysis uses constraints on the I/O behavior of $\Delta$ expressed as (time-domain) IQCs.



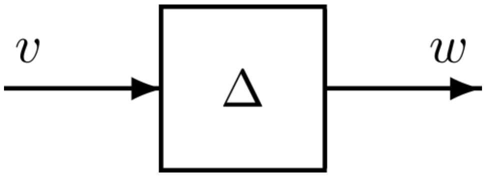

## Definition

$\Delta$ satisfies the IQC on $[0, T]$ defined by a stable filter $\Psi$ and matrix $M$ if:

$$\int_0^T z(t)^T M z(t)\, dt \geq 0 \qquad \forall v \in L_2[0,T] \text{ and } w = \Delta(v)$$


[1] Yakubovich, S-procedure in nonlinear control theory, 1971.
[2] Megretski and Rantzer, System Analysis via Integral Quadratic Constraints, TAC, 1997.

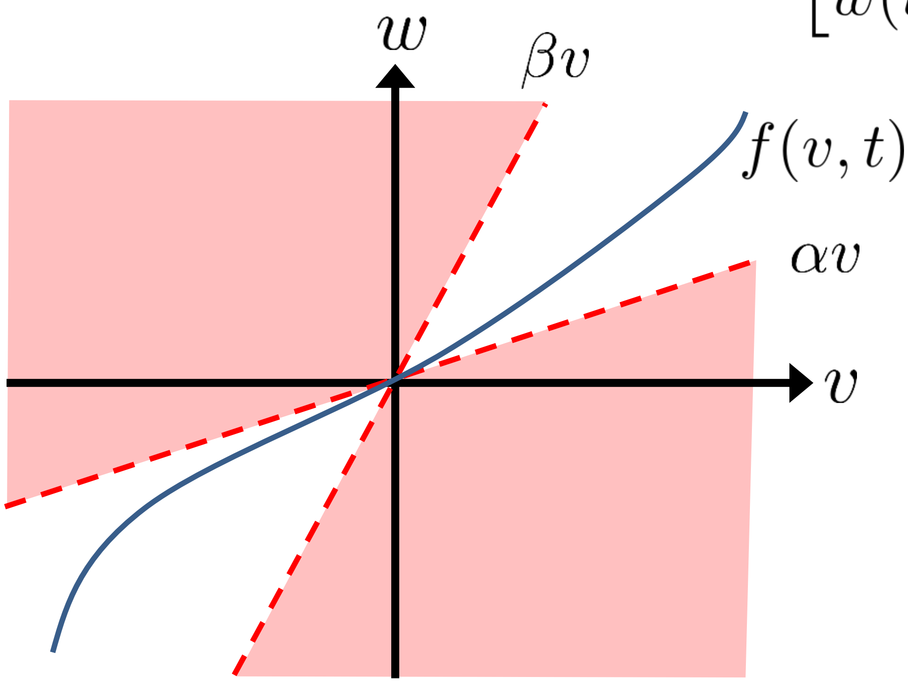
# Example: Sector-bounded Nonlinearity



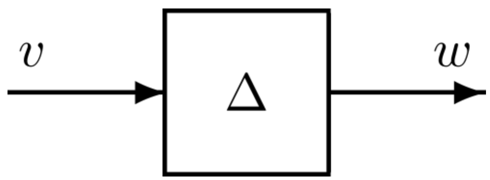

$\Delta$ is a sector-bounded nonlinearity, $f$.

$$(w(t) - \alpha v(t)) \cdot (\beta v(t) - w(t)) \geq 0 \quad \forall t$$

$$\begin{bmatrix} v(t) \\ w(t) \end{bmatrix}^T \begin{bmatrix} -2\alpha\beta & \alpha+\beta \\ \alpha+\beta & -2 \end{bmatrix} \begin{bmatrix} v(t) \\ w(t) \end{bmatrix} \geq 0 \;\; \forall t$$



# Example: Sector-bounded Nonlinearity



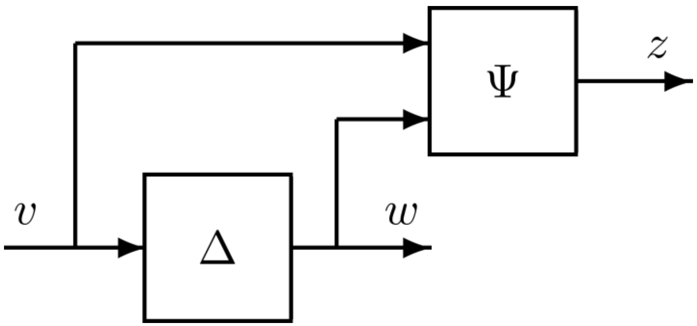

$\Delta$ is a sector-bounded nonlinearity, $f$.

$$(w(t) - \alpha v(t)) \cdot (\beta v(t) - w(t)) \geq 0 \quad \forall t$$

⇕

$$\begin{bmatrix} v(t) \\ w(t) \end{bmatrix}^T \underbrace{\begin{bmatrix} -2\alpha\beta & \alpha+\beta \\ \alpha+\beta & -2 \end{bmatrix}}_{:=M} \underbrace{\begin{bmatrix} v(t) \\ w(t) \end{bmatrix}}_{:=z(t)} \geq 0 \;\; \forall t$$

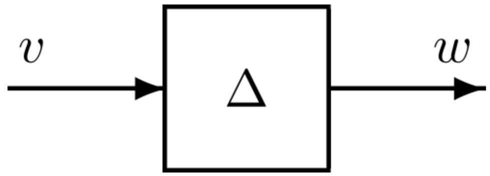
⇓



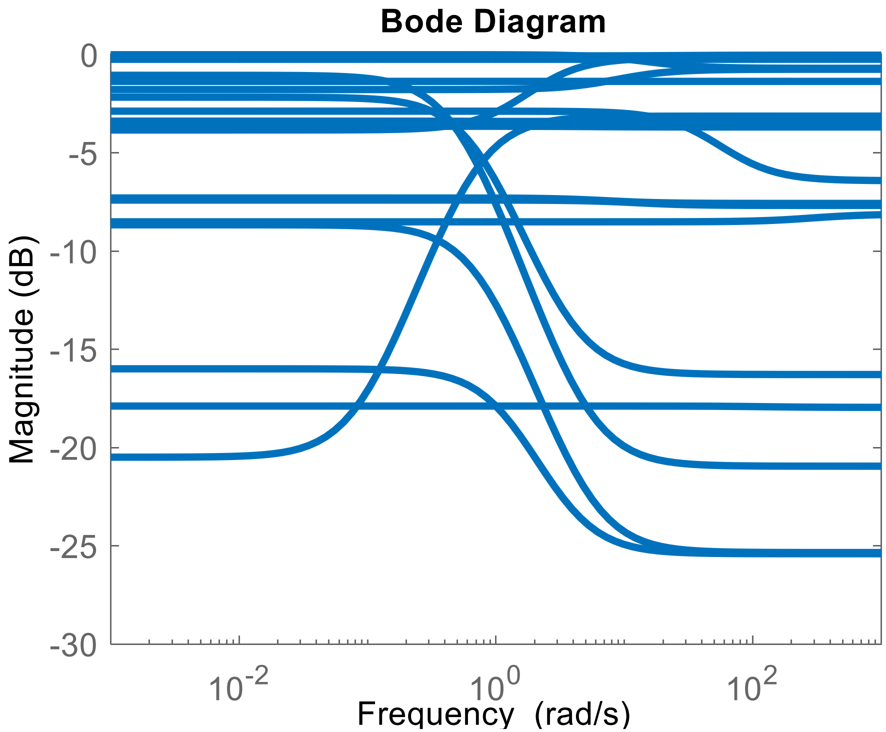

$$\int_0^T z(t)^T M z(t)\, dt \geq 0$$

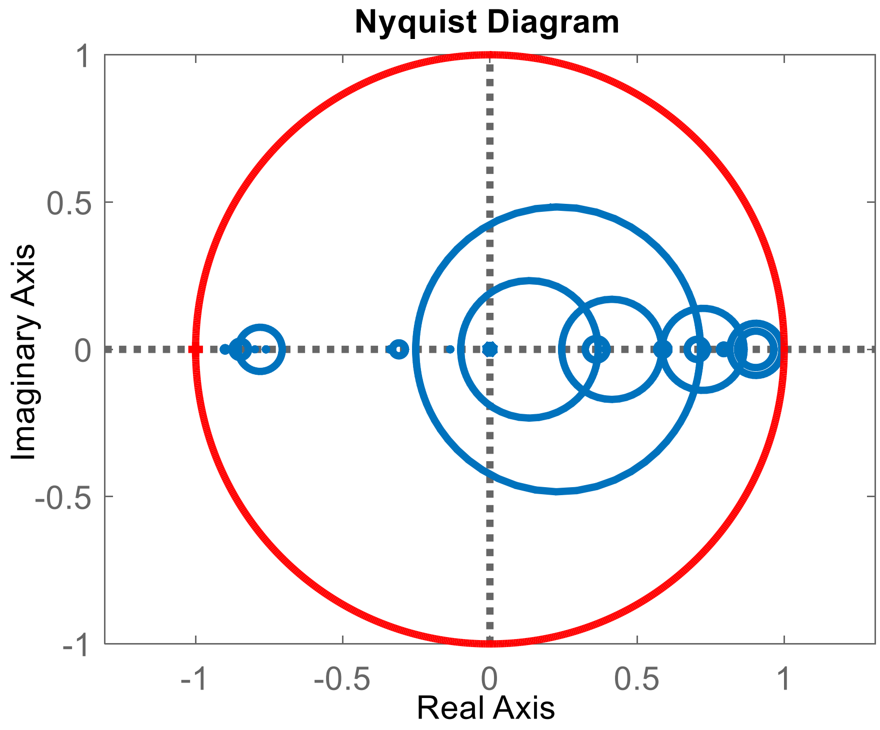
$\Delta$ satisfies the IQC on $[0, T]$ defined by $\Psi := I_2$ and $M$.

# Example: Non-parametric (Dynamic) Uncertainty

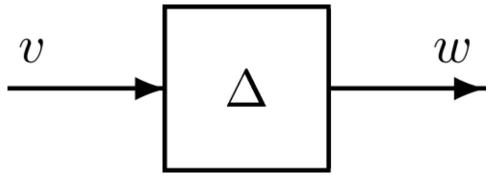



$\Delta$ is stable, LTI with
$$\|\Delta\|_\infty := \sup_\omega |\Delta(\omega)| \leq 1$$



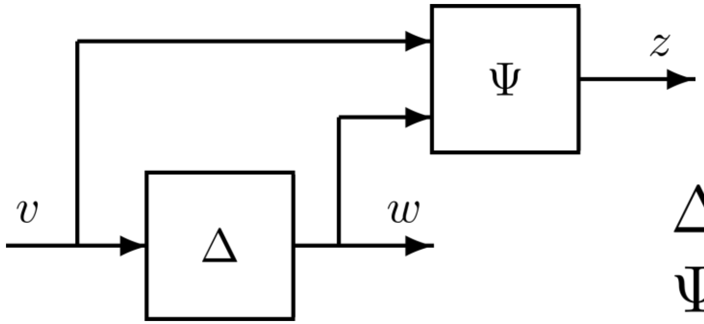



# Example: Non-parametric (Dynamic) Uncertainty [1]




$\Delta$ is stable, LTI with
$\|\Delta\|_\infty := \sup_\omega |\Delta(\omega)| \leq 1$

For any D(s)




$$\int_0^T z(t)^T M z(t)\, dt \geq 0$$

$\Delta$ satisfies the IQC on $[0, T]$ defined by
$\Psi := \text{diag}(D, D)$ and $M := \text{diag}(1, -1)$

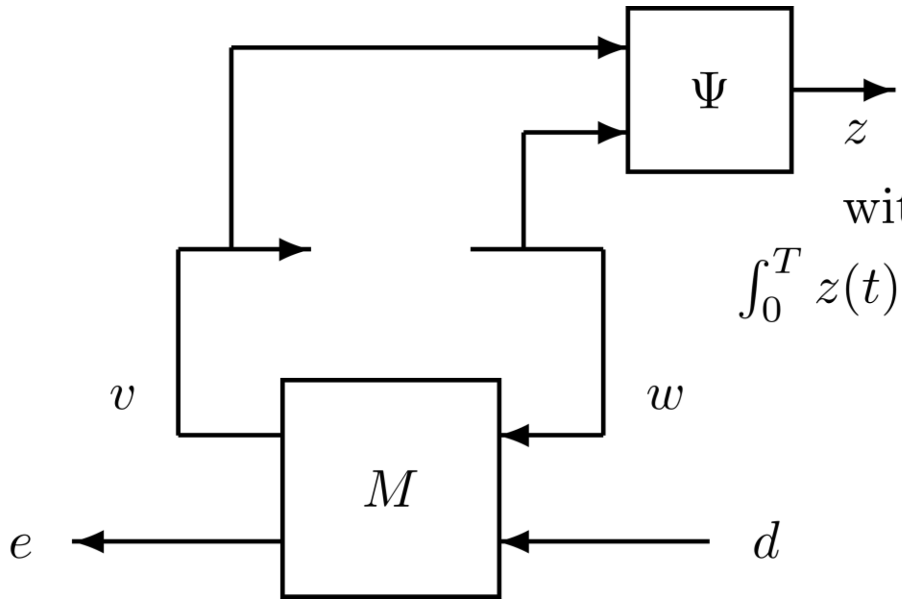

[1] Balakrishnan, Lyapunov Functionals in Complex μ Analysis, TAC, 2002.

# Additional IQC Details

- A dictionary of additional IQC for various uncertainties / nonlinearities is given in [1].
  - IQCs for passive operators, static memoryless nonlinearities (Popov, Zames-Falb), time-delays, real parameters, etc.
  - Many IQCs are specified in the frequency domain
- Most IQCs are related to previous robust stability results
  - IQC for sector nonlinearities related to the circle criterion
  - IQC for LTI uncertainties related to D-scales in $\mu$ analysis
- A technical J-spectral factorization result can be used to convert freq. domain IQCs into time-domain IQCs [2,3].

[1] Megretski & Rantzer, System analysis via IQCs, TAC, 1997. [IQCs derived based on much prior literature]
[2] Seiler, Stability Analysis with Dissipation Inequalities and Integral Quadratic Constraints, TAC, 2015.
[3] Hu, Lacerda, Seiler, Robustness Analysis of Uncertain Discrete-Time System with … IQCs, IJRNC, 2016.

# Robustness Analysis




with $\int_0^T z(t)^T M(t) z(t)\, dt \geq 0$

The robustness analysis is performed on the extended (LTV) system of $(M, \Psi)$ using the constraint on $z$.

$$\begin{bmatrix} \dot{x}_e(t) \\ z(t) \\ e(t) \end{bmatrix} = \begin{bmatrix} \mathcal{A}(t) & \mathcal{B}_1(t) & \mathcal{B}_2(t) \\ \mathcal{C}_1(t) & \mathcal{D}_1(t) & \mathcal{D}_2(t) \\ \mathcal{C}_2(t) & 0 & 0 \end{bmatrix} \begin{bmatrix} x_e(t) \\ w(t) \\ d(t) \end{bmatrix}$$

# Robust Finite Horizon Analysis

## Theorem [1,2]

Assume $\Delta$ satisfies the IQC defined by $(\Psi, M)$.
If there exists $P(\cdot) = P(\cdot)^T$ such that
$(i)$ $P(T) = \mathcal{C}_2(T)^T \mathcal{C}_2(T)$, and
$(ii)$ $V(x,t) := x^T P(t) x$ satisfies

$$\frac{d}{dt} V(x,t) - \gamma^2 d(t)^T d(t) + z(t)^T M z(t) \le 0 \ \ \forall t \in [0,T]$$

then $\|e(T)\|_2 \le \gamma \|d\|_{2,[0,T]}$

## Proof

Integrate dissipation inequality from $t=0$ to $t=T$:

$$\underbrace{V(x(T),T)}_{=e(T)^T e(T)} - \underbrace{V(x(0),0)}_{=0} - \gamma^2 \int_0^T d(t)^T d(t) dt + \underbrace{\int_0^T z(t)^T M z(t) dt}_{\ge 0} \le 0$$


[1] Moore, Finite Horizon Robustness Analysis using IQCs, MS Thesis, Berkeley, 2015.
[2] Seiler, Moore, Meissen, Arcak, Packard, Finite Horizon Robustness Analysis of LTV Systems Using IQCs, arXiv 2018 and Automatica 2019.

# Robust Finite Horizon Analysis

## Theorem [1,2]

Assume $\Delta$ satisfies the IQC defined by $(\Psi, M)$.
If there exists $P(\cdot) = P(\cdot)^T$ such that
$(i)$ $P(T) = \mathcal{C}_2(T)^T \mathcal{C}_2(T)$, and
$(ii)$ $V(x,t) := x^T P(t) x$ satisfies

$$\frac{d}{dt} V(x,t) - \gamma^2 d(t)^T d(t) + z(t)^T M z(t) \leq 0 \;\; \forall t \in [0,T]$$

then $\|e(T)\|_2 \leq \gamma \|d\|_{2,[0,T]}$

Dissipation inequality can be recast as a differential LMI:

$$\begin{bmatrix} \dot{P} + \mathcal{A}^T P + P\mathcal{A} & P\mathcal{B}_1 & \mathcal{B}_2 \\ \mathcal{B}_1^T P & 0 & 0 \\ \mathcal{B}_2^T P & 0 & -\gamma^2 I \end{bmatrix} + (\cdot)^T M \begin{bmatrix} \mathcal{C}_1 & \mathcal{D}_1 & \mathcal{D}_2 \end{bmatrix} \preceq 0$$

$\forall t \in [0,T]$

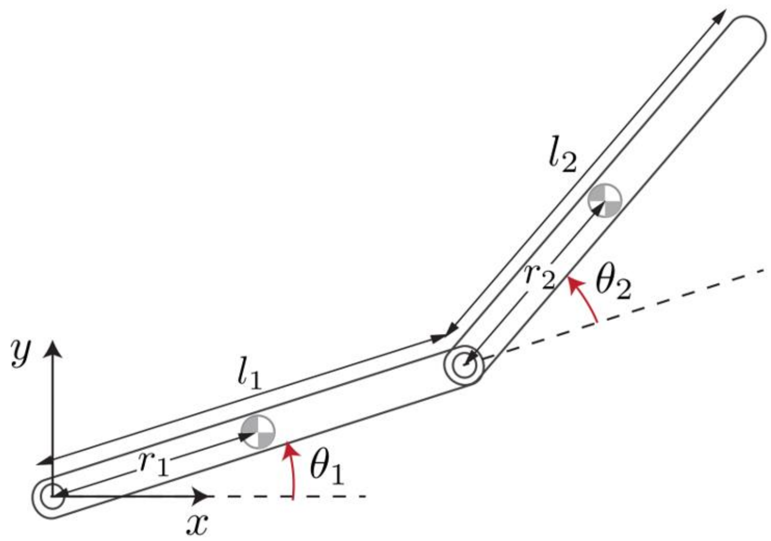
[1] Moore, Finite Horizon Robustness Analysis using IQCs, MS Thesis, Berkeley, 2015.
[2] Seiler, Moore, Meissen, Arcak, Packard, Finite Horizon Robustness Analysis of LTV Systems Using IQCs, arXiv 2018 and Automatica 2019.

# Numerical Algorithms and Software

- **Robustness Algorithms**
  - Differential LMI can be "solved" via convex optimization using basis functions for $P(\cdot)$ and gridding on time [1].
  - A more efficient algorithm mixes the differential LMI and a related Riccati Differential Equation condition [2].
  - Similar methods developed for LPV [4,5] and periodic systems [6].
- **LTVTools Software [3]**
  - Time-varying state space system objects, e.g. obtained from Simulink snapshot linearizations.
  - Includes functions for nominal and robustness analyses.


[1] Moore, Finite Horizon Robustness Analysis using IQCs, MS Thesis, Berkeley, 2015.
[2] Seiler, Moore, Meissen, Arcak, Packard, Finite Horizon Robustness Analysis of LTV Systems Using IQCs, arXiv 2018 and Automatica 2019.
[3] https://z.umn.edu/LTVTools
[4] Pfifer & Seiler, Less Conservative Robustness Analysis of LPV Systems Using IQCs, IJRNC, 2016.
[5] Hjartarson, Packard, Seiler, LPVTools: A Toolbox for Modeling, Analysis, & Synthesis of LPV Systems, 2015.
[6] Fry, Farhood, Seiler, IQC-based robustness analysis of discrete-time LTV systems, IJRNC 2017.

# Two-Link Robot Arm




Two-Link Diagram [1]

Nonlinear dynamics [MZS]:

$$\dot{\eta} = f(\eta, \tau, d)$$

where

$$\eta = [\theta_1, \dot{\theta}_1, \theta_2, \dot{\theta}_2]^T$$
$$\tau = [\tau_1,\ \tau_2\ ]^T$$
$$d = [d_1,\ d_2\ ]^T$$

$\tau$ and $d$ are control torques and disturbances at the link joints.


[1] R. Murray, Z. Li, and S. Sastry. *A Mathematical Introduction to Robot Manipulation*, 1994.

# Nominal Trajectory in Cartesian Coordinates

# Analysis

Nonlinear dynamics:

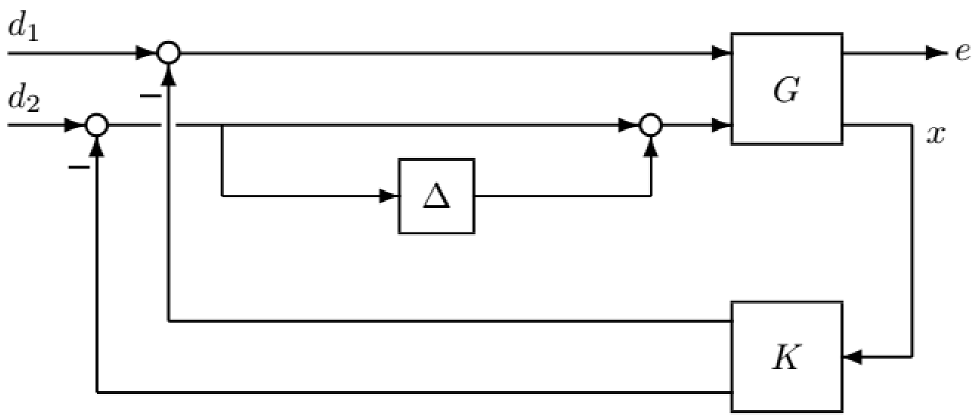
$$\dot{\eta} = f(\eta, \tau, d)$$

Linearize along the finite –horizon trajectory $(\bar{\eta}, \bar{\tau}, d = 0)$

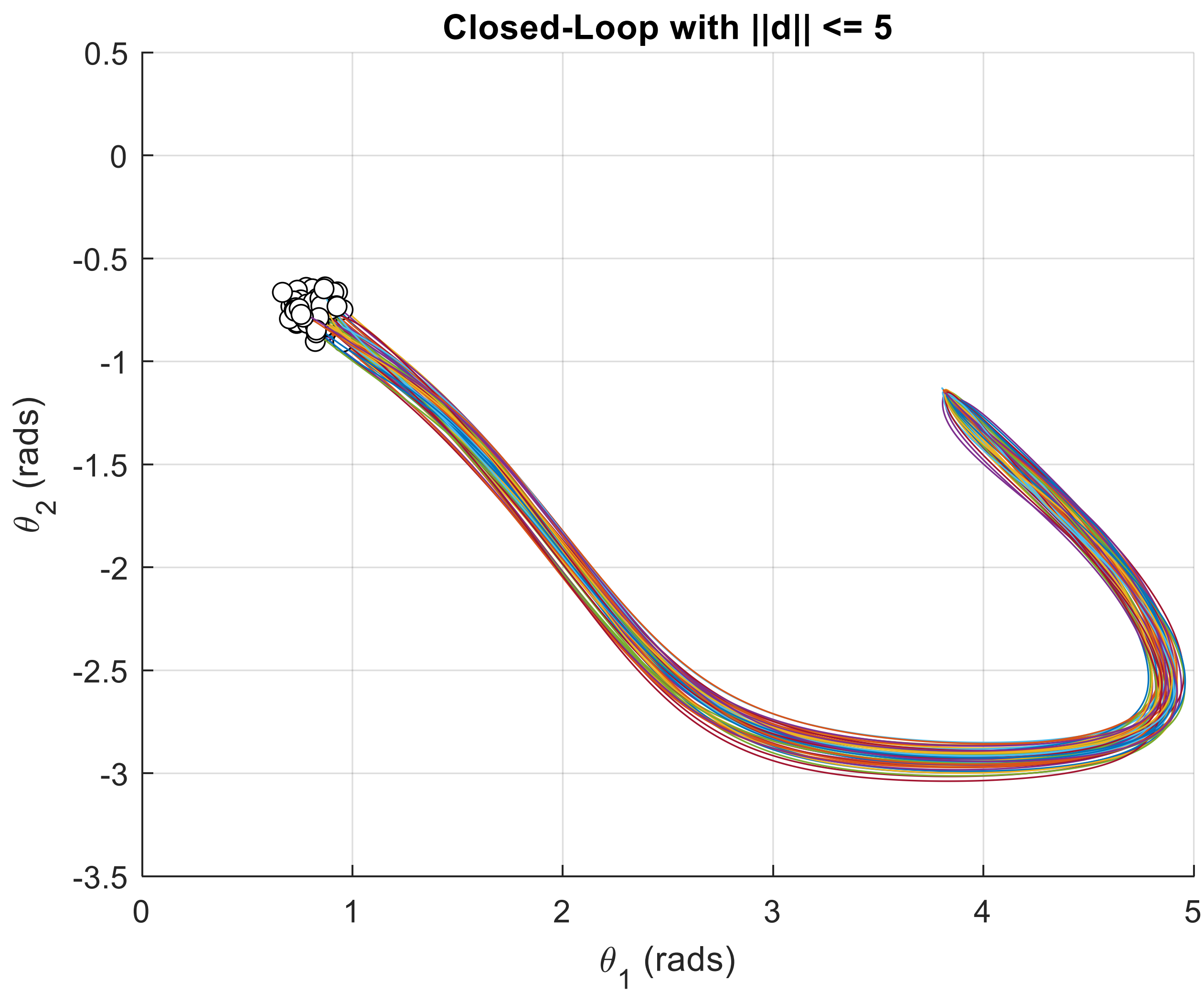
$$\dot{x} = A(t)x + B(t)u + B(t)d$$

Design finite-horizon state-feedback LQR gain.

Goal: Compute bound on the final position accounting for disturbances and LTI uncertainty $\Delta$ at 2nd joint.

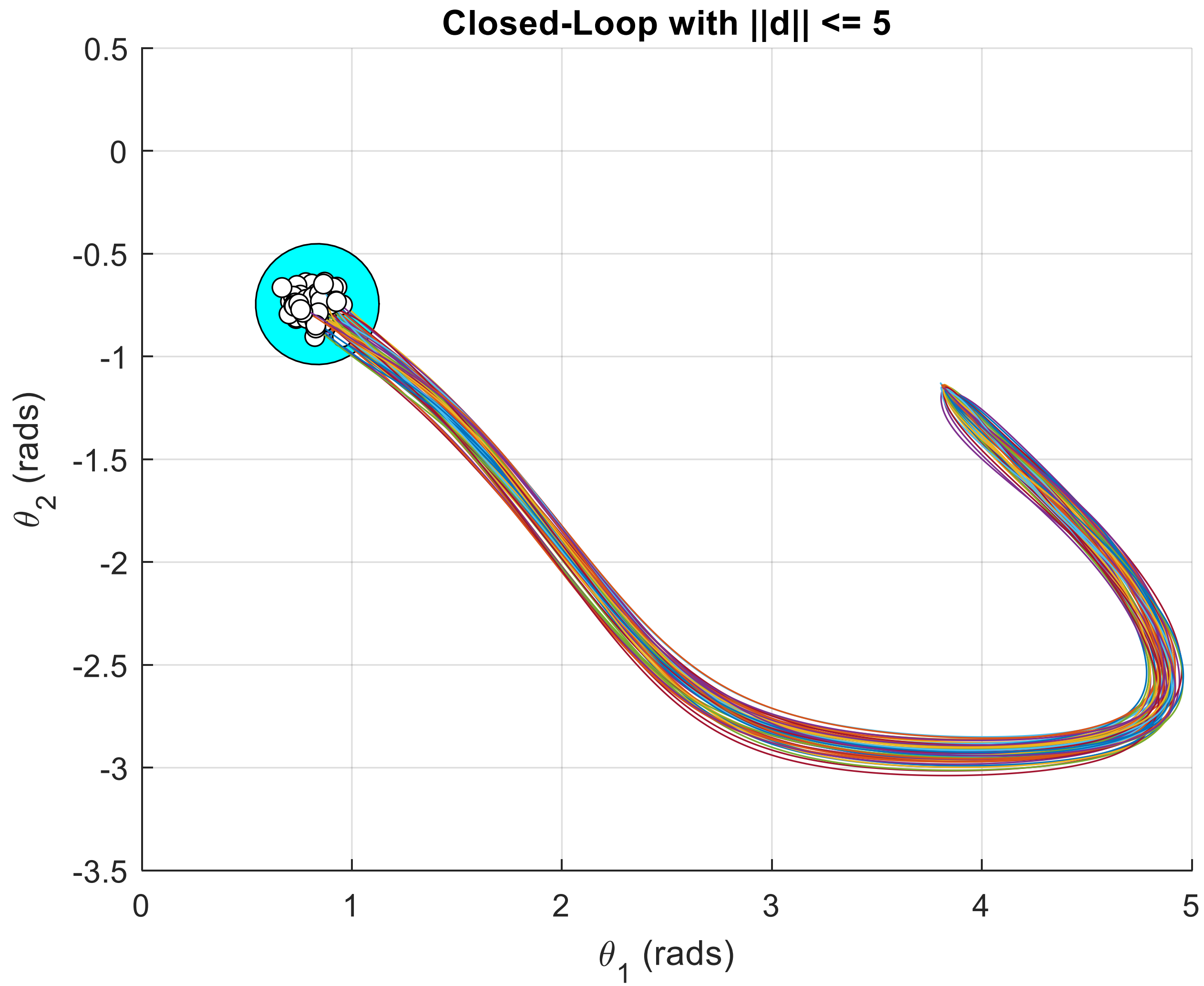


# Monte-Carlo Simulations




LTV simulations with randomly sampled disturbances and uncertainties (overlaid on nominal trajectory).

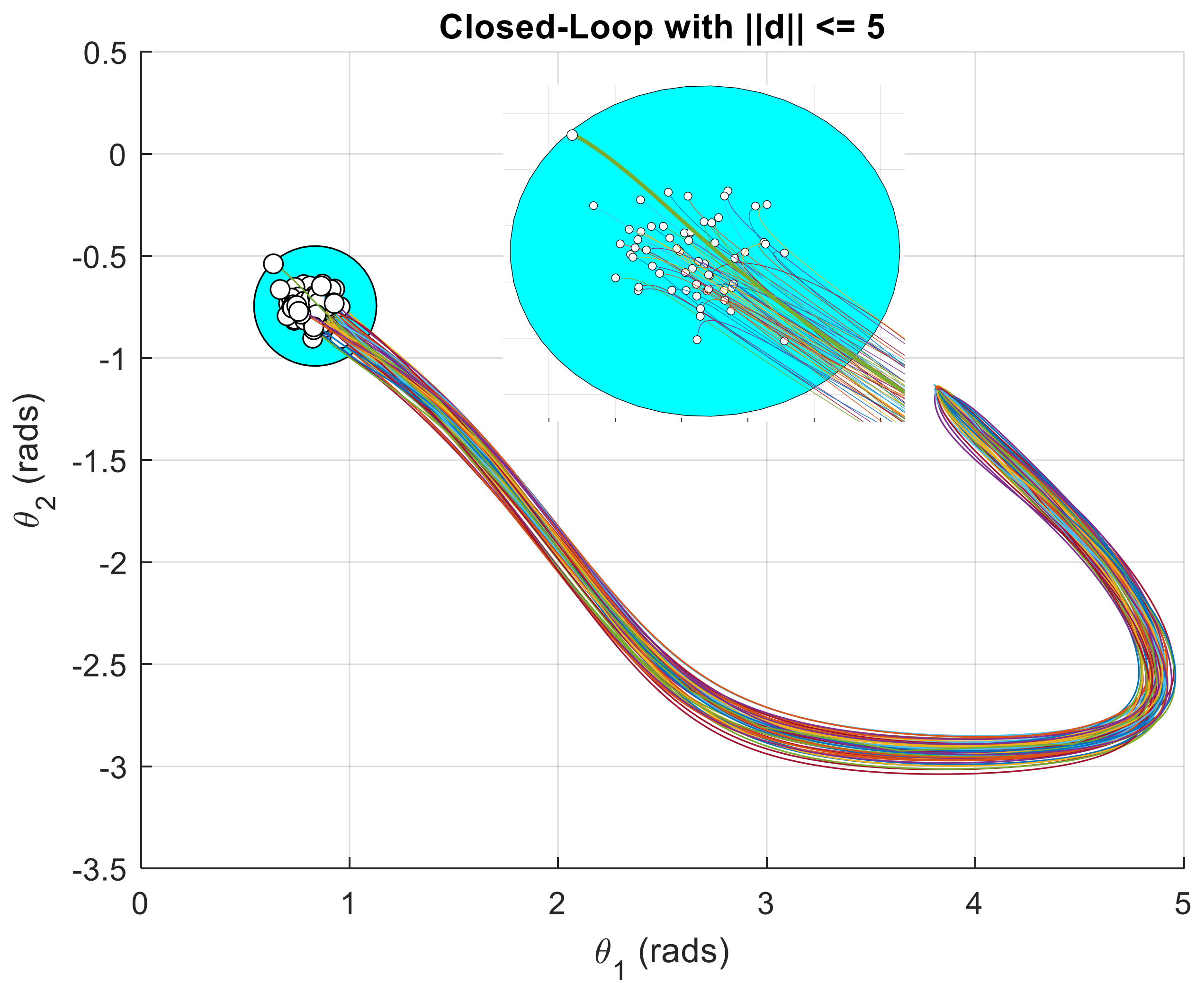
# Robustness Bound




Cyan disk is bound computed in 102 sec using IQC/DI method

Bound accounts for disturbances $\|d\| \leq 5$ and $\|\Delta\| \leq 0.8$

# Worst-Case Uncertainty / Disturbance




Randomly sample Δ to find "bad" perturbation and compute corresponding worst-case disturbance using method in [1].

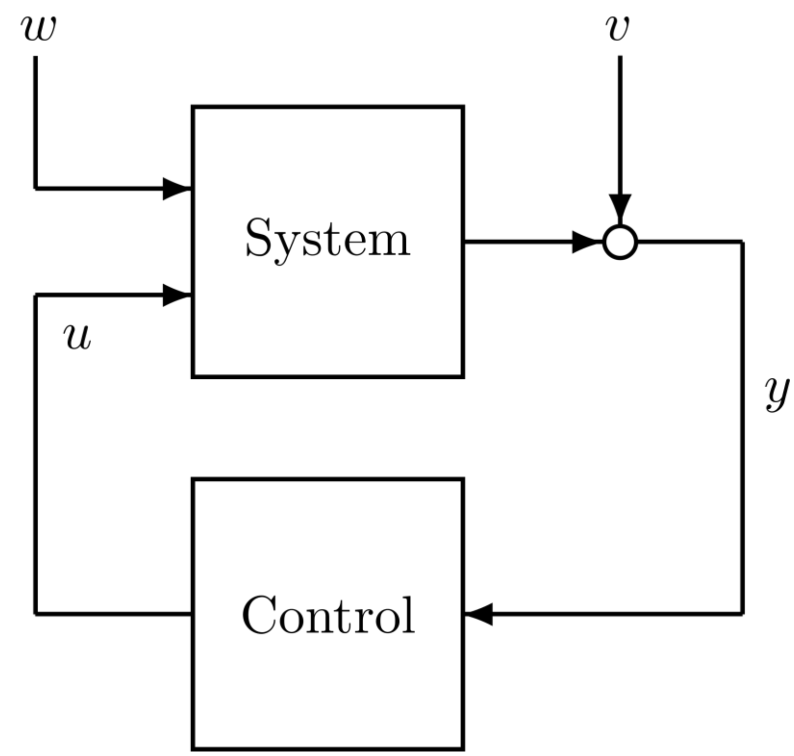

[1] Iannelli, Seiler, Marcos, Construction of worst-case disturbances for LTV systems…, 2019.

# Outline

- Brief Overview of Robust Control
- Robustness of Time-Varying Systems
- **Future Directions**
  - **Robustness in Reinforcement Learning**
  - Optimization as Robust Control
- Conclusions

# “Model-Free” Reinforcement Learning

- **Goal:** Train a control policy from data to maximize a cumulative reward
  - Training data obtained from a simulator or the real system
  - Often assume state feedback
  - Many algorithms (Q-learning, value iteration, policy iteration, policy search) [1,2,3]
  - Algorithms have close connections to dynamic programming and optimal control.





[1] D.P. Bertsekas, “Reinforcement Learning and Optimal Control,” 2019.

[2] R.S. Sutton and A.G. Barto, “Reinforcement Learning: An Introduction,” 2018.

[3] C. Szepesvári, “Algorithms for Reinforcement Learning,” 2010.

# Is Robustness an Issue in RL?

**Training via simulation**

**Training on real system**

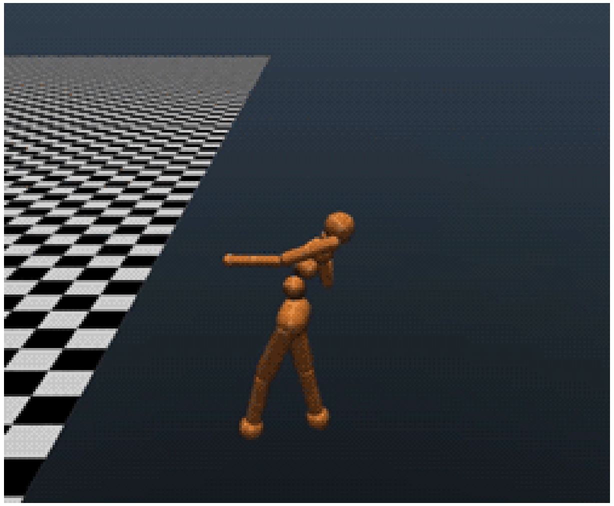
# Is Robustness an Issue in RL?

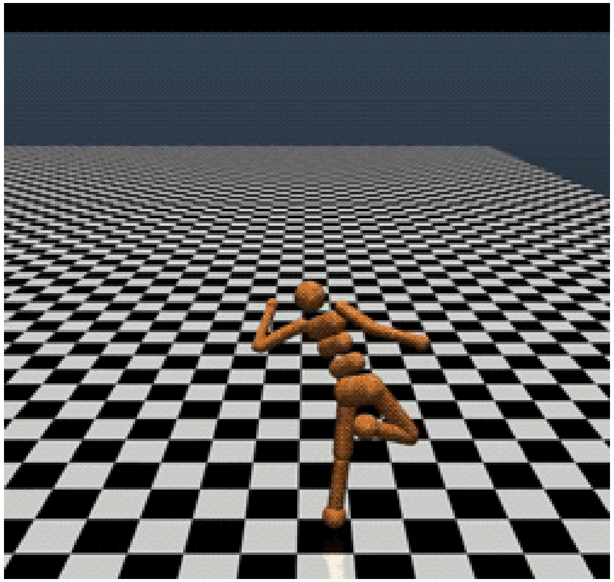
**Training via simulation**

- Training can exploit flaws in the simulator [1].
- Loss of performance transitioning from simulator to real system.

**Training on real system**

Robotic Walking in MuJoCo





[1] B. Recht, "A Tour of Reinforcement Learning," arXiv, 2018.

# Is Robustness an Issue in RL?

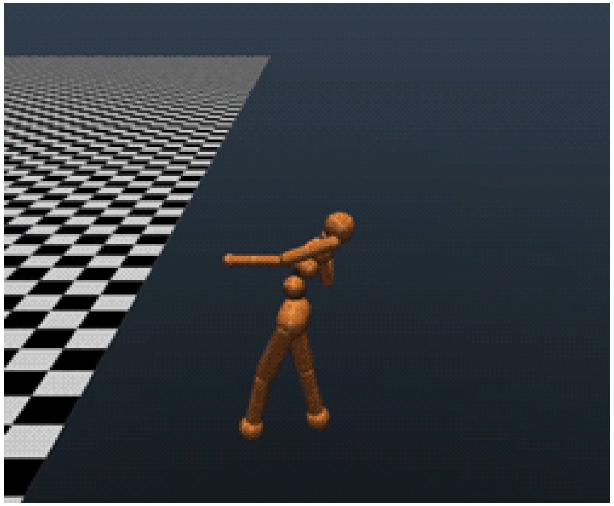
**Training via simulation**

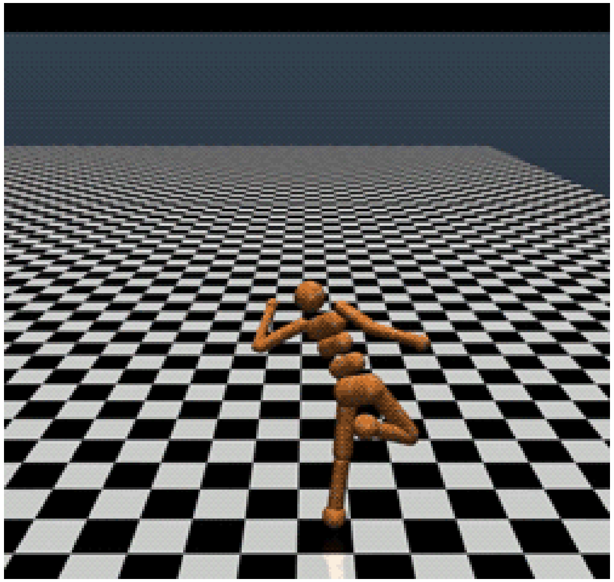
- Training can exploit flaws in the simulator [1].
- Loss of performance transitioning from simulator to real system.

**Training on real system**

- Part to part variation (train on one system and implement on many)
- Changes in system dynamics over time (temperature dependence, environmental effects, etc.…)

Robotic Walking in MuJoCo





[1] B. Recht, "A Tour of Reinforcement Learning," arXiv, 2018.

# Initial Investigations [1]

- Use linear optimal control problems to understand performance of RL techniques
  - RL provides most benefit for problems that can’t be addressed by standard system ID + linear optimal control
  - However, LTI problems can be used as “test” cases

- Develop (model-free) methods to recover robustness
  - Model uncertainty is different from process noise
  - **What is the appropriate regularizer?**

[1] Venkataraman & Seiler, Recovering Robustness in Model-Free Reinforcement Learning, ’18 arXiv and ‘19 ACC (accepted).

# Linear Quadratic Gaussian (LQG)

Minimize $$J_{LQG}(u) := \lim_{N\to\infty} \frac{1}{N} E\left[\sum_{t=0}^{N} x_t^T Q x_t + u_t^T R u_t\right]$$

Subject To: $$\begin{aligned} x_{t+1} &= A x_t + B u_t + B_w w_t \\ y_t &= C x_t + v_t \end{aligned}$$

The optimal controller has an observer/state-feedback form

$$\begin{aligned} \hat{x}_{t+1} &= A\hat{x}_t + B u_t + L\left(y_t - C\hat{x}_t\right) \\ u_t &= -K\hat{x}_t \end{aligned}$$

Gains *(K,L)* computed by solving two Riccati equations.

This solution is model-based, i.e. it uses data *A,B,C,* etc

# Reinforcement Learning

- Partially Observable Markov Decision Processes (POMDPs)
  - Set of states, *S*
  - Set of actions, *A*
  - Reward function, $r: S \times A \to \mathbb{R}$
  - State transition probability, *T*
  - Set of observations and observation probability, *O*
- Many methods to synthesize a control policy from input/output data to maximize the cumulative reward

$$J_{RL}(a) := E\left[\sum_{t=0}^{N} r(s_t, a_t)\right]$$

- The LQG problem is a special case of this RL formulation

# Doyle's Example ('78 TAC)

- LQR state-feedback regulators have provably good margins.
- Doyle's example shows that LQG (output-feedback) regulators can have arbitrarily small input margins.

**Honeywell** Interoffice Correspondence

Date: August 23, 1977

To: C. A. Harvey

From: J. C. Doyle

Location: S&RC, Research

cc: L. Q. Gaussian
J. A. Hauge
A. P. Kizilos
A. F. Konar
E. E. Yore
N. R. Zagalsky
Systems and Control Technology

Subject: "Guaranteed Margins for LQG Regulators"

A B S T R A C T

There aren't any.

All engineers who have been using LQG methodology may pick up their Nichols charts from the supply room.

# Doyle's Example ('78 TAC)

- LQR state-feedback regulators have provably good margins.
- Doyle's example shows that LQG regulators can have arbitrarily small input margins.
- Doyle's example can also be solved within RL framework using direct policy search:

$$\begin{aligned} z_{t+1} &= A_K(\theta) z_t + B_K(\theta) y_t \\ u_t &= C_K(\theta) z_t \end{aligned}$$

where

$$A_K(\theta) := \begin{bmatrix} 0 & \theta_1 \\ 1 & \theta_2 \end{bmatrix}, B_K(\theta) := \begin{bmatrix} 1 \\ 0 \end{bmatrix}, C_K^T(\theta) := \begin{bmatrix} \theta_3 \\ \theta_4 \end{bmatrix}$$

- RL will converge to the optimal LQG control with infinite data collection. Thus RL can also have poor margins.

# What is an Appropriate Regularizer for Robustness?

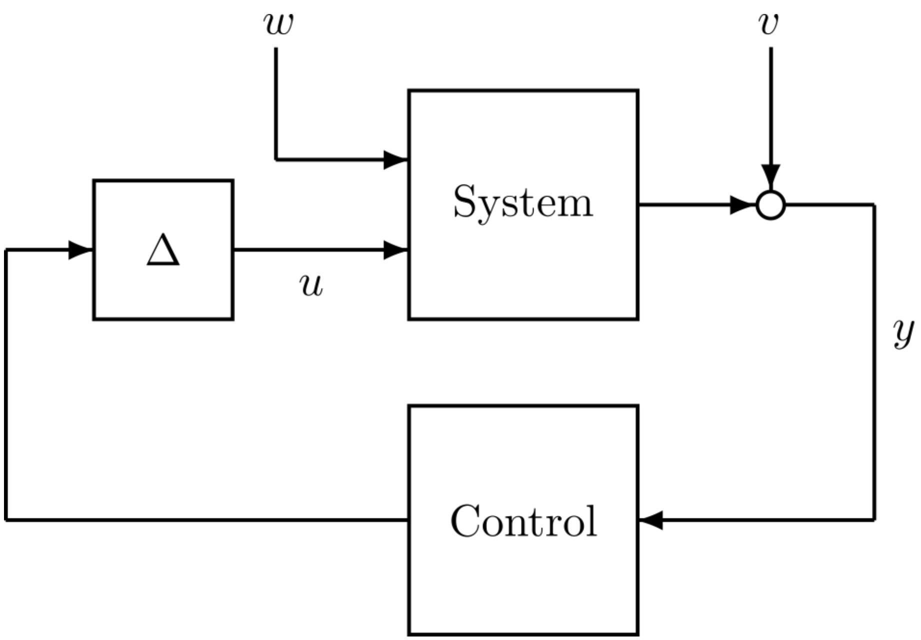
- Increase process noise during training?
  - This causes margins to decrease on Doyle's example
  - **Process noise is not model uncertainty**
- Modify reward to increase state penalty or decrease control penalty?
  - Again, this causes margins to decrease on Doyle's example
  - Trading performance vs. robustness via the reward function can be difficult or counter-intuitive

# Proposed Method to Recover Robustness

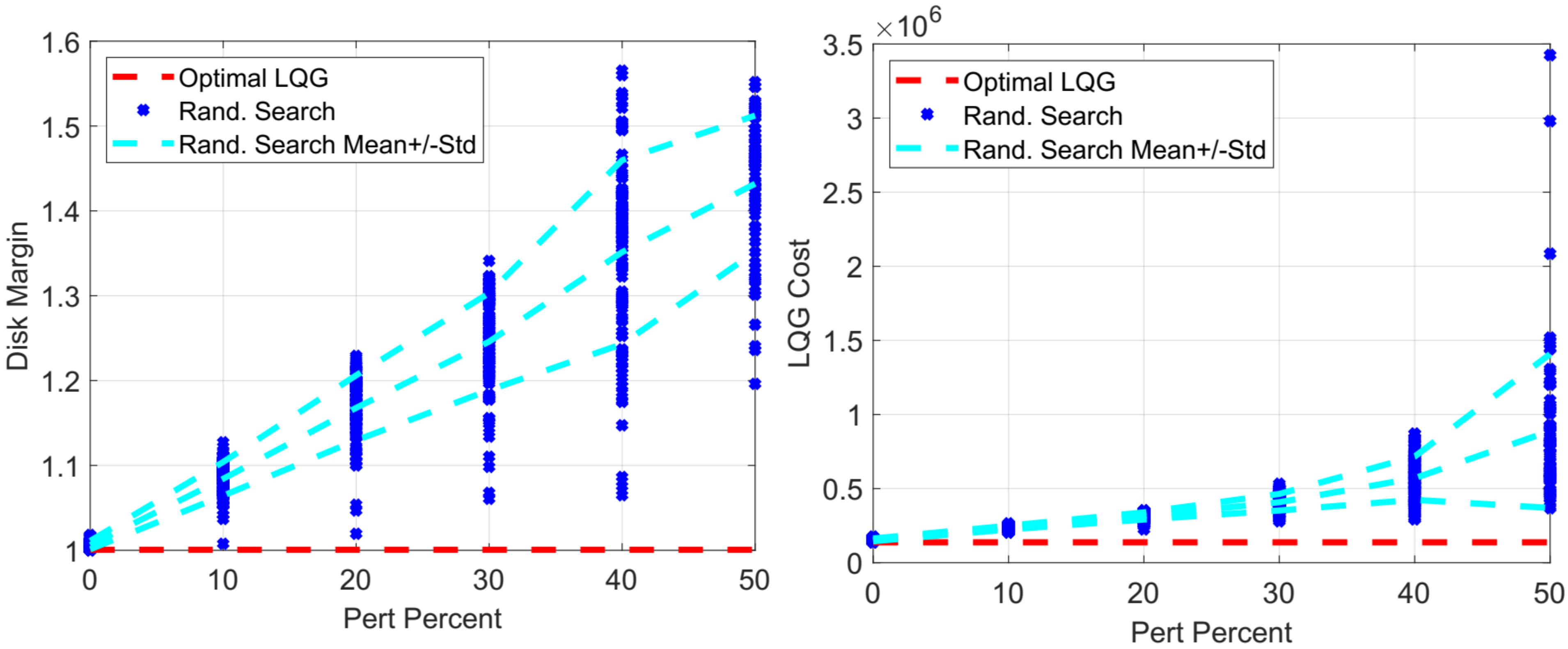
Inject synthetic gain/phase variations at the plant input (and output?) during the training phase.

$\Delta = 1 + \delta$ where

$\delta$ is *U[-b,b]*

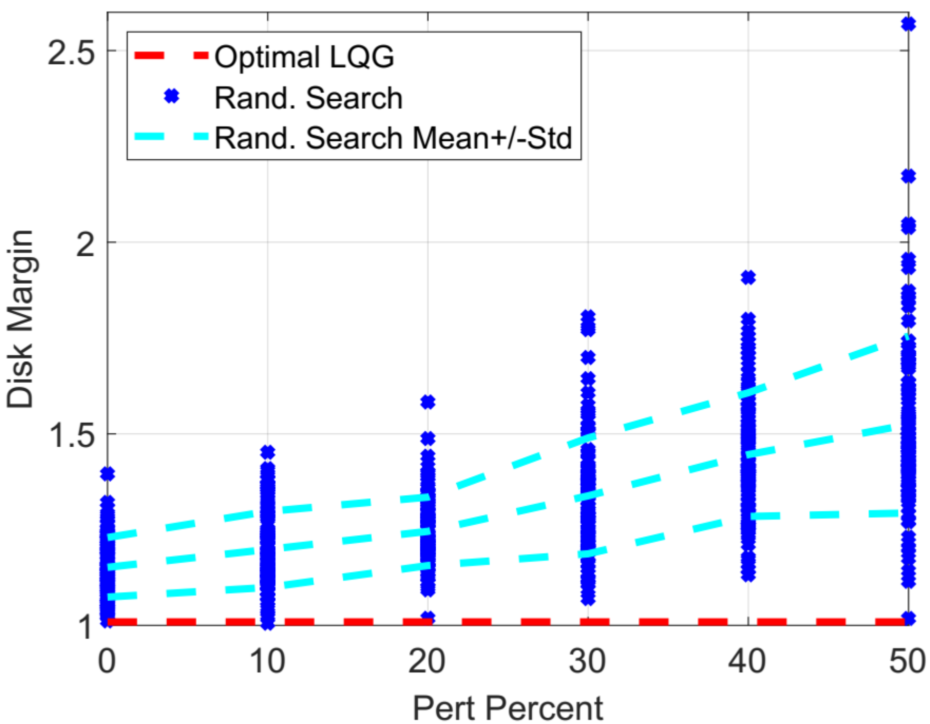


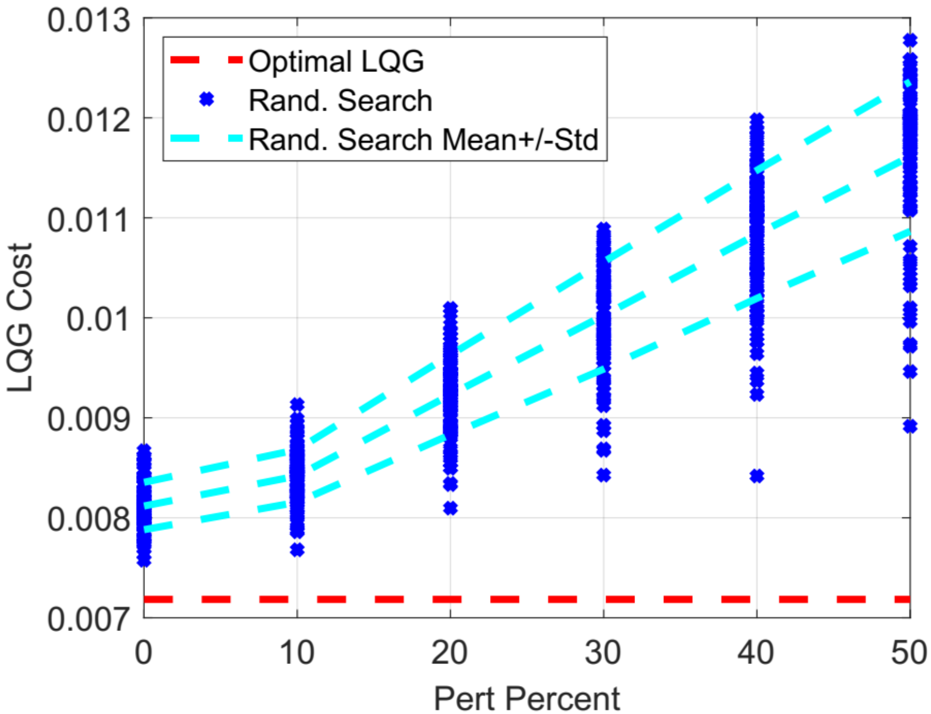
# Results On Doyle's Example



# Results on Simplified Flexible System

- Model has 4-states (Rigid body and lightly damped modes)
- RL applied to 3-state controller parameterization
  - LQG controller is not in the control policy parameterization
  - Still converges to policy with small margins
  - Robustness recovered with synthetic perturbations during training



# Longer Term Goals/Questions

- Develop (model-free) methods to recover robustness. What is the appropriate regularizer?
- Understand how to merge lower level (model-based) control with higher level (model-free) methods.
  - What is an appropriate merging point?
  - What is a useful model abstraction for higher level (model-free) methods?
- Can we make any rigorous claims about the proposed method? Performance certification?
- What are fundamental performance limits on RL policies?
  - What will an RL-trained algorithm do for a fundamentally difficult problem, e.g. $G(s) = (s-1)/(s-2)$?

# Outline

- Brief Overview of Robust Control
- Robustness of Time-Varying Systems
- **Future Directions**
  - Robustness in Reinforcement Learning
  - **Optimization as Robust Control**
- Conclusions

# First-order Optimization Algorithms

Assumptions on $f$

- Strongly convex ($m$)
- Lipschitz gradients ($L$)

$$\min_{x \in \mathbf{R}^n} f(x)$$

<u>Gradient Descent</u>

$$x_{k+1} = x_k - \alpha \nabla f(x_k)$$

First-Order Algorithm

- Input: Gradient at iterate
- Output: Next iterate

<u>Heavy-Ball</u>

$$x_{k+1} = x_k - \alpha \nabla f(x_k) + \beta(x_k - x_{k-1})$$

<u>Nesterov's Method</u>

$$x_{k+1} = y_k - \alpha \nabla f(y_k)$$
$$y_k = (1+\beta)x_k - \beta x_{k-1})$$

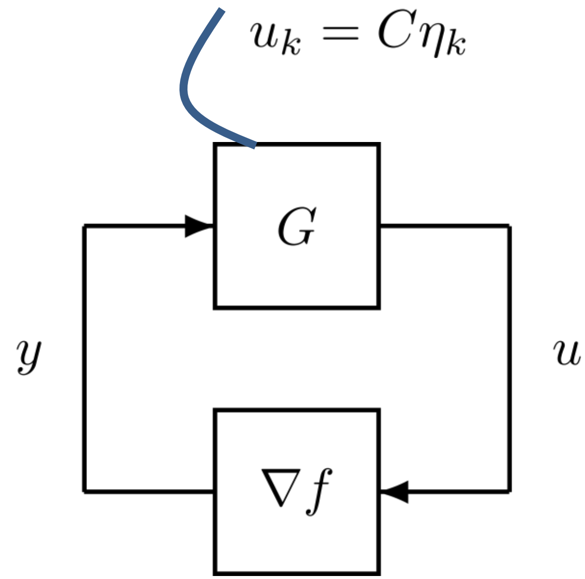
# First-order Optimization as Robust Control [1]

$$\min_{x \in \mathbf{R}^n} f(x)$$

## Robust Control Perspective

- Uncertain plant, $\nabla f$
- Controller $G$ (algorithm) is finite-dim, strictly proper, LTI system

$$\eta_{k+1} = A\eta_k + By_k$$
$$u_k = C\eta_k$$




## Automated Analysis with IQC/SDP

- Characterize $\nabla f$ with IQCs
- “Small LMIs” to certify convergence rate
- Analytical proofs guided by SDP solns.
- Extensions including algorithm design

[1] Lessard, Recht, Packard, Analysis and Design of Optimization Algorithms via IQCs, SIAM, 2015

# Extension to Stochastic Optimization

## Finite Sum Minimization

$$\min_{x\in\mathbf{R}^n} \frac{1}{n}\sum_{i=1}^{n} f^i(x)$$

- Certain convexity/Lipschitz assumptions
- Application to empirical risk minimization,

  e.g. in supervised learning.

# Extension to Stochastic Optimization

## Finite Sum Minimization

- Certain convexity/Lipschitz assumptions
- Application to empirical risk minimization

$$\min_{x \in \mathbf{R}^n} \frac{1}{n} \sum_{i=1}^{n} f^i(x)$$

## Stochastic Gradient is widely used

- Fixed stepsize: Convergence to tolerance of optimal
- Decreasing stepsize: Sublinear convergence

## Many recent methods (SAGA, Finito, SDCA) with linear convergence and similar iteration cost as SG.

<u>SAGA</u>

Randomly sample $i_k$ at each step

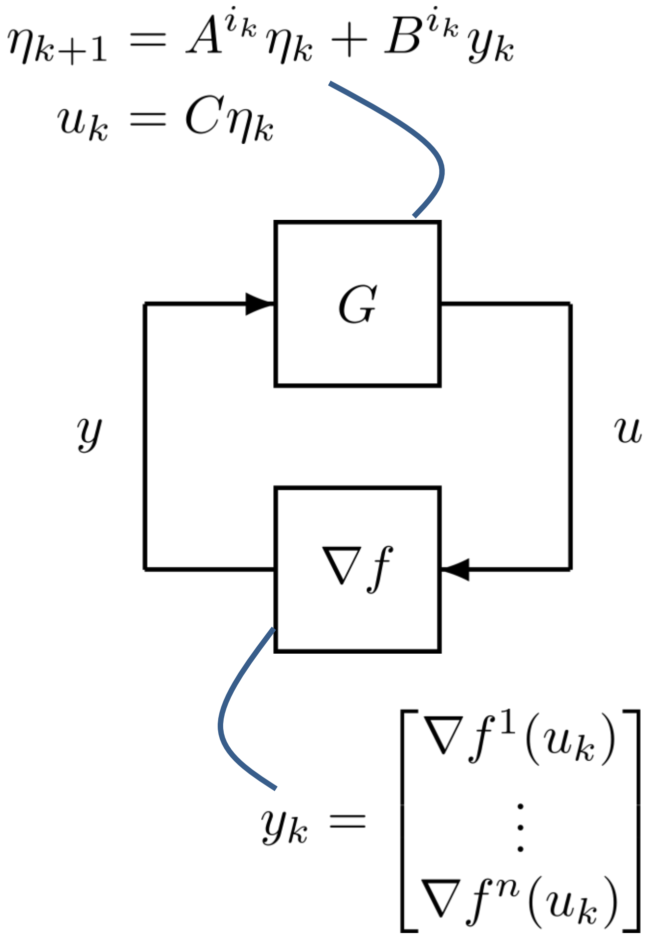
$$x_{k+1} = x_k - \alpha \left( \nabla f^{i_k}(x_k) - y_k^{i_k} + \tfrac{1}{n} \textstyle\sum_{i=1}^{n} y_k^i \right)$$

$$y_{k+1}^i := \begin{cases} \nabla f^i(x_k) & \text{if } i = i_k \\ y_k^i & \text{else} \end{cases}$$

# Extension to Stochastic Optimization [1,2]

Express stochastic optimization techniques with:

- Uncertain plant, $\nabla f$
- Markov Jump System representation for optimization algorithm




Automated Analysis with IQC/SDP

- Characterize $\nabla f$ with IQCs
- "Small" SDPs to certify convergence-rate
- Analytical proofs guided by SDP solns.


[1] Hu, Seiler, Rantzer, A Unified Analysis of Stochastic Optimization Methods Using Jump System Theory and Quadratic Constraints, COLT 2017.

[2] Hu, A Robust Control Perspective on Optimization of Strongly-Convex Functions, Ph.D. , 2016.

# Longer Term Goals/Questions

- Determine if finite horizon analysis analysis tools can be used to assess convergence rates.
  - Related work on finite horizon Performance Estimation Problem (PEP) [1]
- Are IQC rate bounds tight for strictly convex, Lipschitz bounded functions? If no, then for what class are they tight?
  - Initial results prove tightness for stability boundary but likely not true in general [2].
- Can methods from robust synthesis be used to design algorithms with faster convergence?
  - IQC synthesis is non-convex so this would require some heuristics
  - Would require new robust synthesis methods for jump systems.

[1] Taylor, Hendrickx, Glineur, Smooth Strongly Convex Interpolation and Exact Worst-case Performance of First-order Methods, Math. Prog., 2017.

[2] Badithela and Seiler, Analysis of the Heavy-ball algorithm using IQCs, accepted to ACC, 2019.

[3] Van Scoy, Freeman, Lynch, The Fastest Known Globally Convergent First-Order Method for Minimizing Strongly Convex Functions, IEEE CSL, 2018.

# Outline

- Brief Overview of Robust Control
- Robustness of Time-Varying Systems
- Future Directions
- **Conclusions**

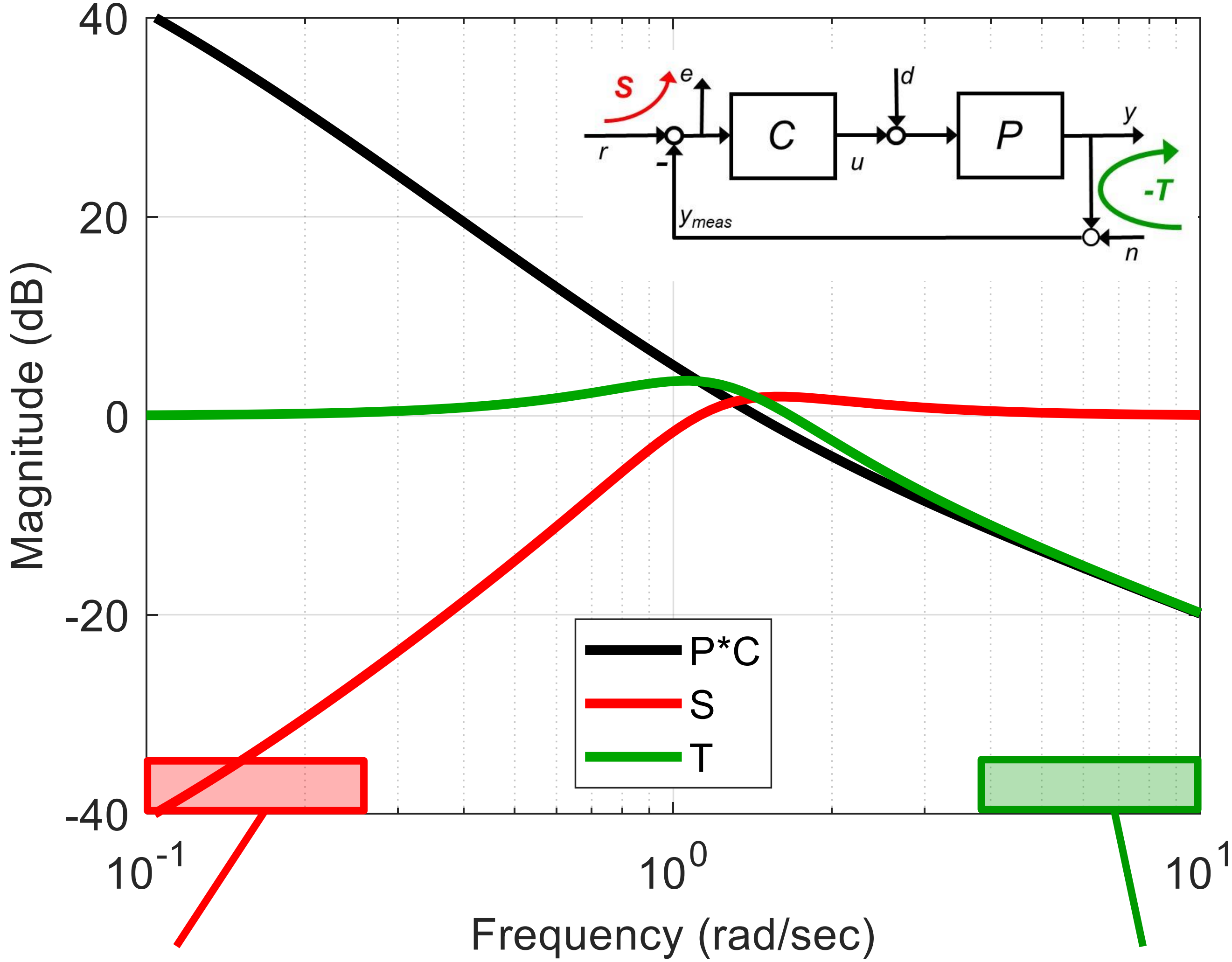
# Conclusions

- Robust control has a long history with many successes
  1. Multivariable Optimal Control
  2. Fundamental Limitations of Dynamics & Control
  3. Uncertainty Modeling and Robustness Analysis
- Robust control techniques can solve emerging problems
  1. Robustness in controls designed via data-driven (RL) methods
  2. Optimization as Robust Control
- Acknowledgements:
  
  - Funding: NSF, AFOSR, ONR, NASA, Seagate, MSI, Xcel RDF, MnDrive
  - Past PhDs & Visitors: Annoni, Hu, Honda, Kotikalpudi, Lacerda, Ossmann, Peni, Pfifer, Takarics, Theis, Venkataraman, Wang
  

https://www.aem.umn.edu/~SeilerControl/

# Backup Slides

# Typical S+T=1 Tradeoff




$$P = \frac{1}{s+0.1}$$

$$C = \frac{s+1.5}{s}$$

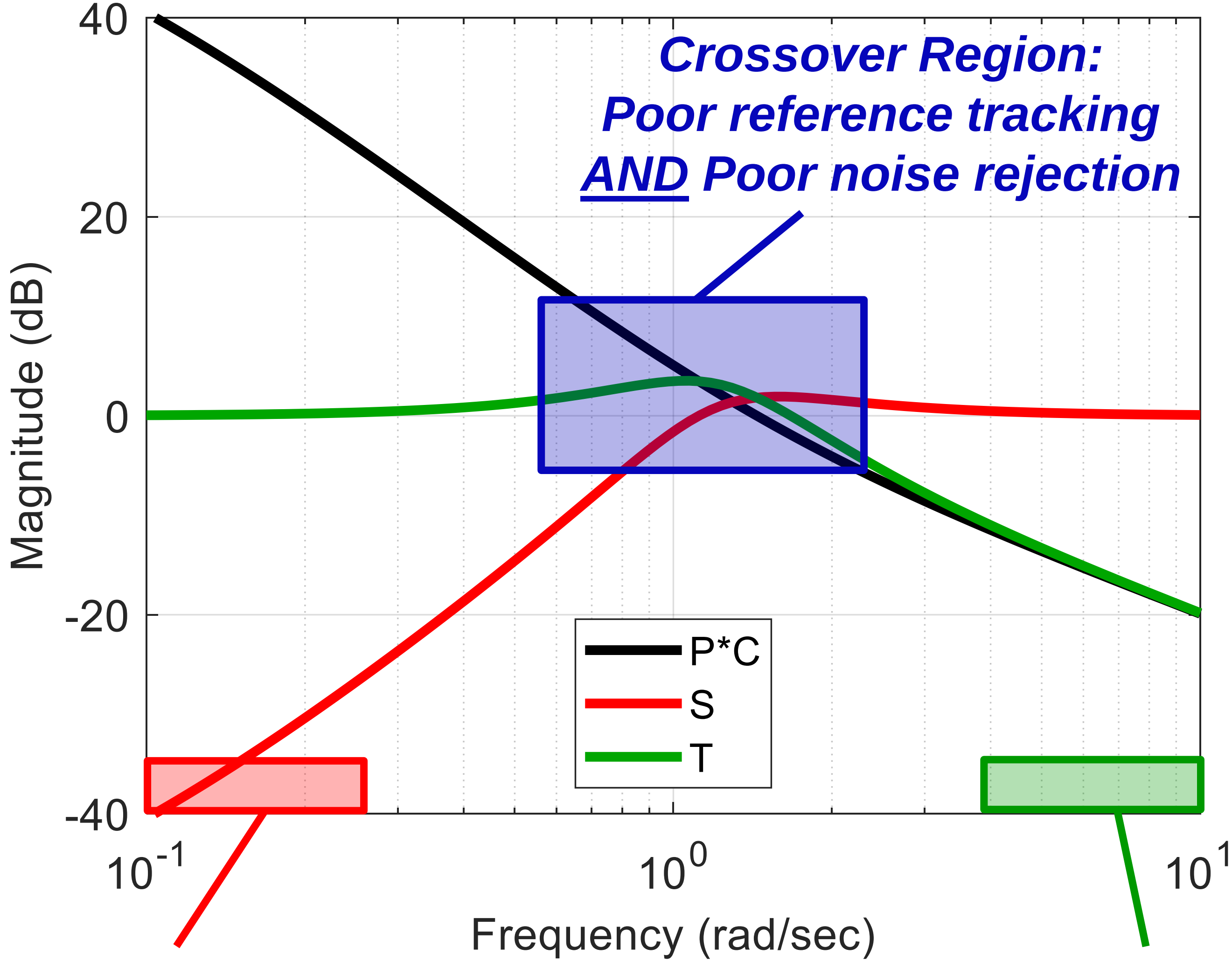
$$S = \frac{1}{1+PC}$$

$$T = \frac{PC}{1+PC}$$

***Large loop gain |PC|:***
***Good reference tracking***
***Poor noise rejection***

***Small loop gain |PC|:***
***Poor reference tracking***
***Good noise rejection***

# Typical S+T=1 Tradeoff




$$P = \frac{1}{s+0.1}$$

$$C = \frac{s+1.5}{s}$$

$$S = \frac{1}{1+PC}$$

$$T = \frac{PC}{1+PC}$$

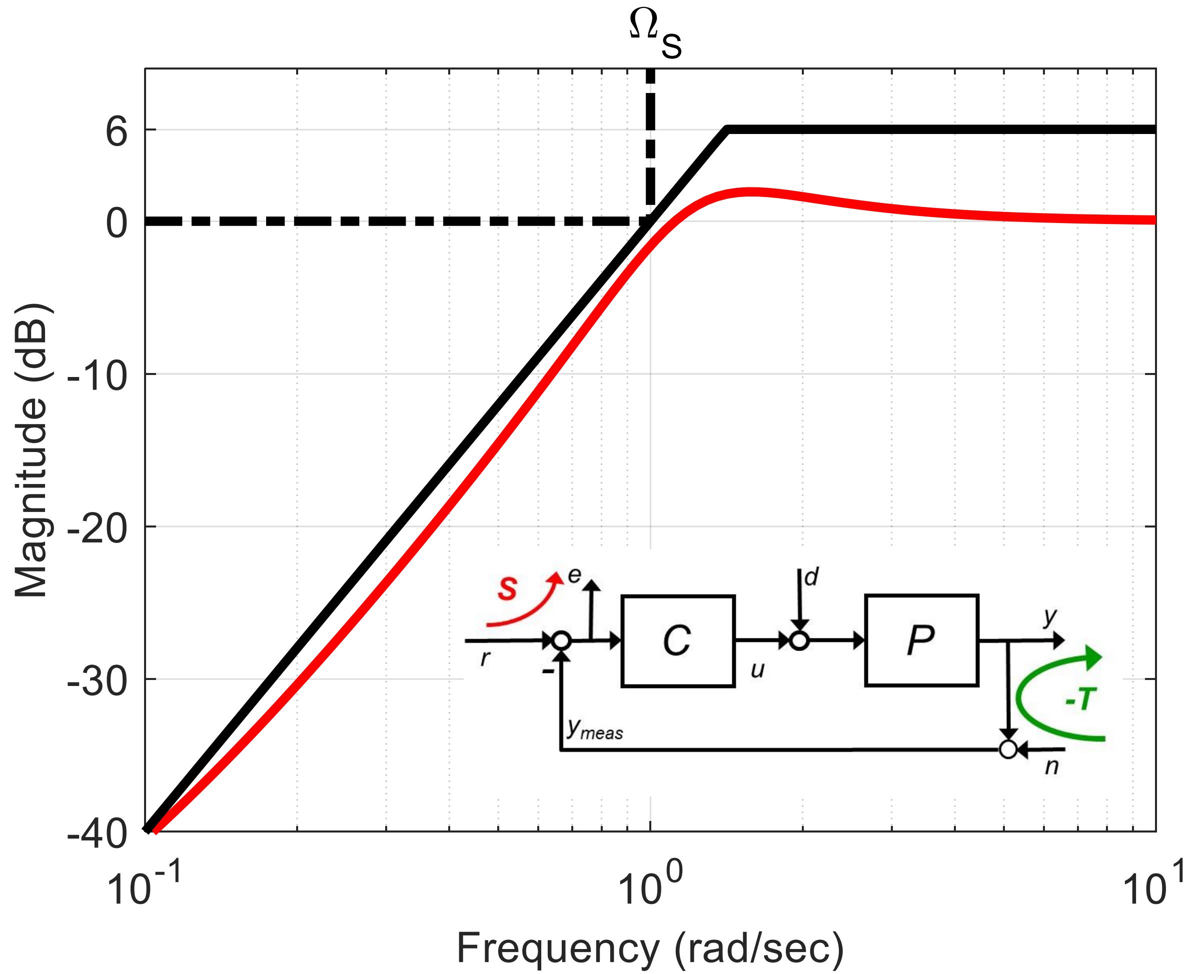
***Large loop gain |PC|:***
***Good reference tracking***
***Poor noise rejection***

***Small loop gain |PC|:***
***Poor reference tracking***
***Good noise rejection***

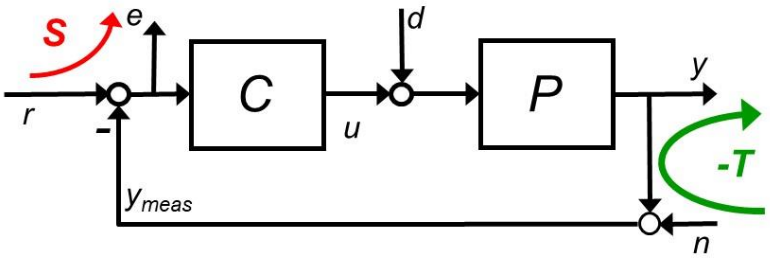
# Typical Sensitivity Objectives

- **Performance:** “Small” |S| up to 0 dB bandwidth $\Omega_S$
- **Robustness:** |S|≤ 2 (=6dB) at all frequencies (No Peaks)

Typical sensitivity response (red) and design objectives (black)



# Bode Integral Theorem [1,2]




Assume *PC* has relative degree 2 and *S(s)* is stable. Then:

$$\int_0^{\infty} \ln |S(j\omega)| \, d\omega = \pi \sum_{k=1}^{N_u} \mathrm{Re}(p_k) \geq 0$$

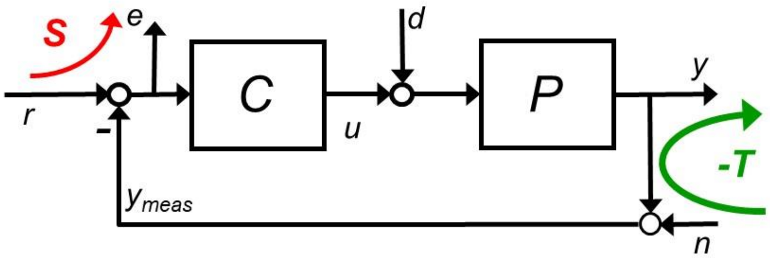
where $p_k$ are the unstable (RHP) poles of *PC*.

(Note: $|S| \text{ (dB)} \approx 8.7 \ln |S|$ )

[1] Bode, Network Analysis and Feedback Amplifier Design, 1945.
[2] Freudenberg and Looze, Frequency Domain Properties of Scalar and Multivariable Feedback Systems, Springer-Verlag, 1988.

# Bode Integral Theorem [1,2]

Assume *PC* has relative degree 2 and *S(s)* is stable. Then:




$$\int_0^{\infty} \ln |S(j\omega)| \, d\omega = \pi \sum_{k=1}^{N_u} \mathrm{Re}(p_k) \geq 0$$

where $p_k$ are the unstable (RHP) poles of *PC*.

(Note: $|S| \ (\mathrm{dB}) \approx 8.7 \ln |S|$ )

**This a key conserved quantity in feedback design.** Improving performance (e.g. increased bandwidth) comes at the expense of reduced robustness (peak in |*S*|) [3].


[1] Bode, Network Analysis and Feedback Amplifier Design, 1945.
[2] Freudenberg and Looze, Frequency Domain Properties of Scalar and Multivariable Feedback Systems, Springer-Verlag, 1988.
[3] Stein, Respect the Unstable, Bode Lecture, 1989 (and IEEE CSM, 2003)

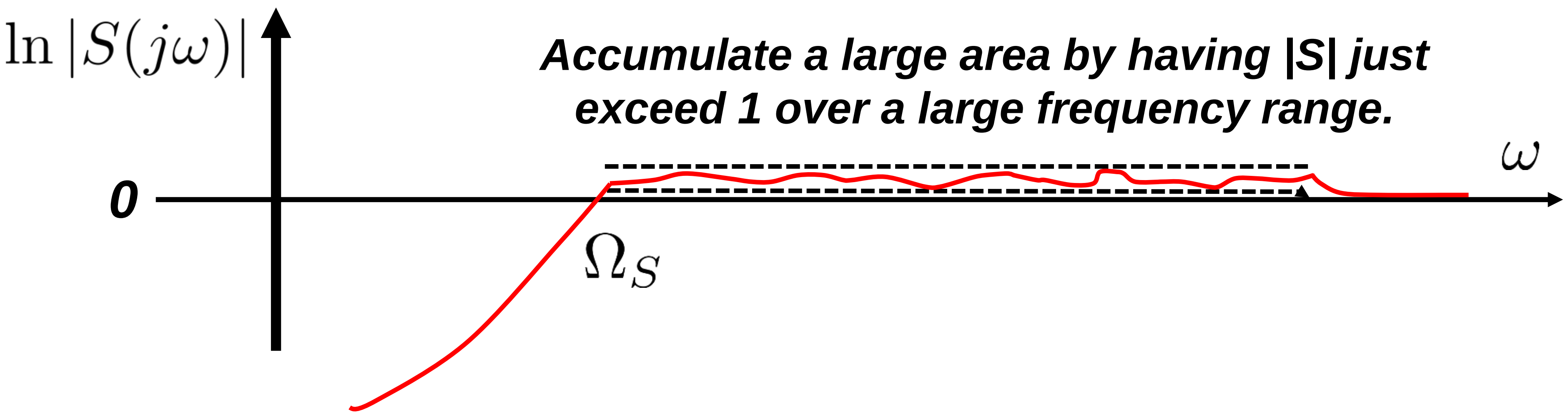
# Bode Integral Theorem and "Peaking"

A procedure to avoid peaking could be:

- Obtain significant Sensitivity reduction over $[0, \Omega_S]$.

  This incurs a large negative integral which must be balanced.
- Maintain $|S(j\omega)|$ slightly larger than 1 over a wide interval.

  This incurs a positive integral balancing the negative integral.
- Make *|PC|* approach 0 quickly at higher frequencies so that |S| quickly approaches 1.

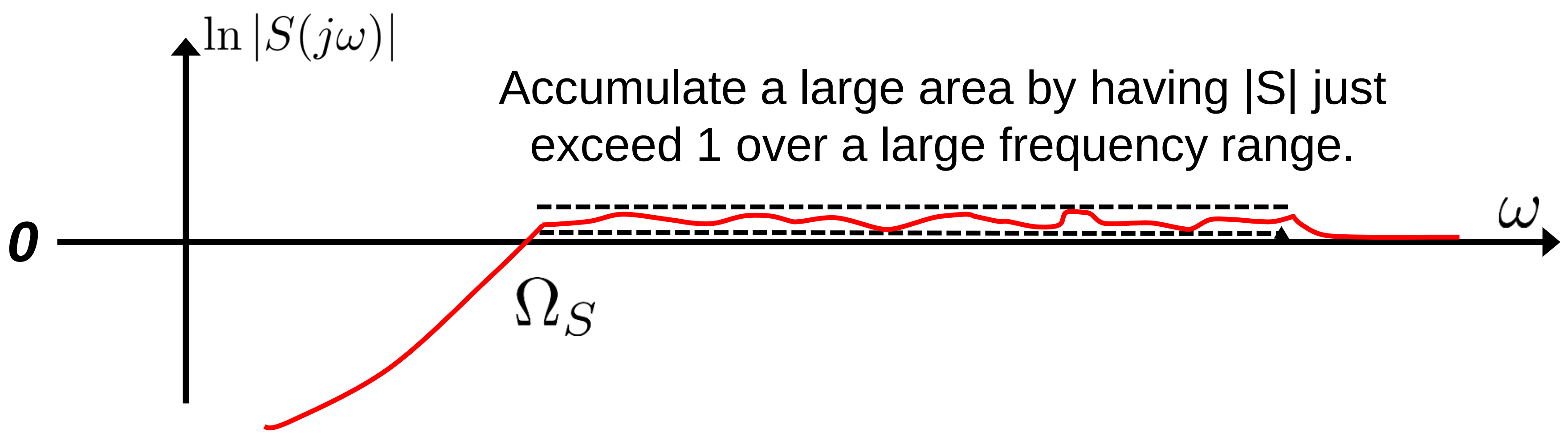


# Available Bandwidth

The Bode Integral theorem may appear to be a minor constraint, e.g. spreading area over a large frequency band.

Stein ('89 Bode Lecture, '03 CSM):

*a key fact about physical systems is that they do not exhibit good frequency response fidelity beyond a certain bandwidth. … Let us call that bandwidth the "available bandwidth,"* $\Omega_a$



# Available Bandwidth

The Bode Integral theorem may appear to be a minor constraint, e.g. spreading area over a large frequency band.

Stein (‘89 Bode Lecture, ’03 CSM):

*a key fact about physical systems is that they do not exhibit good frequency response fidelity beyond a certain bandwidth. … Let us call that bandwidth the “available bandwidth,”* $\Omega_a$

The available bandwidth due to physical (hardware) constraints requires positive area be accumulated over a finite frequency band.

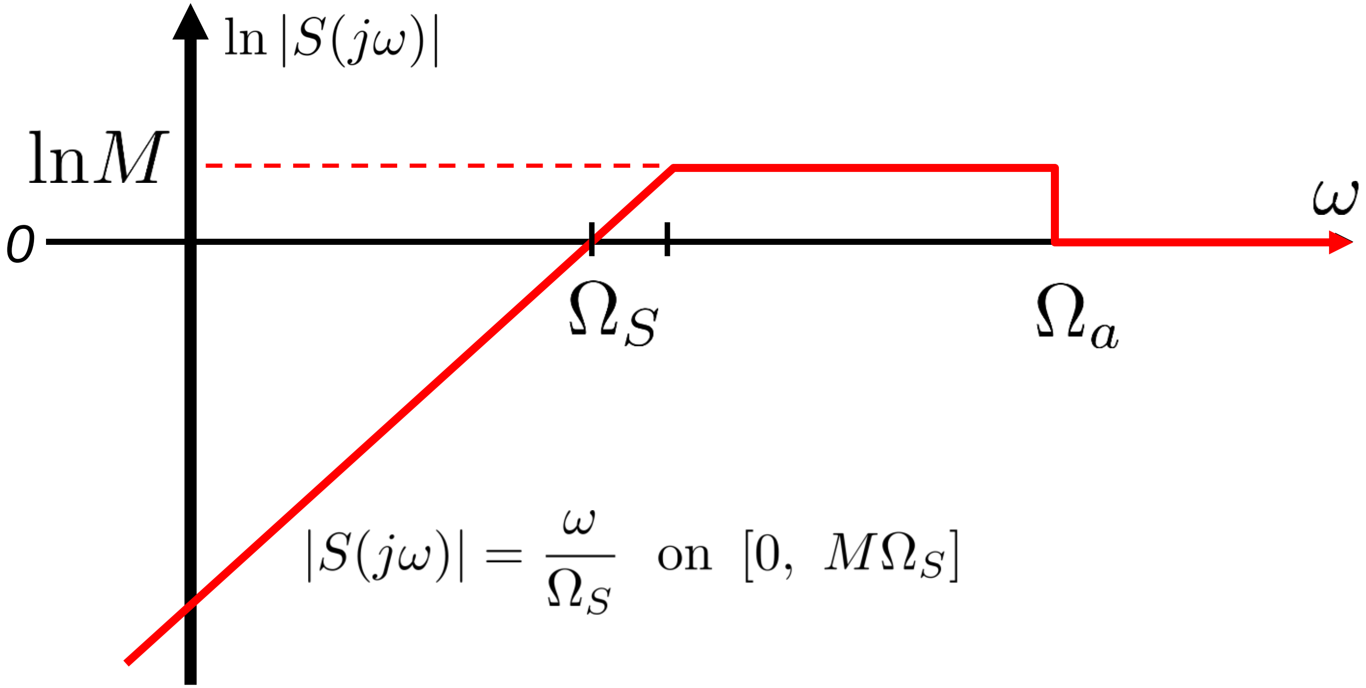
**Consequence: Improving performance (e.g. increased bandwidth) comes at the expense of reduced robustness (peak in $|S|$).**

# Consequence of Available Bandwidth

$|PC|$ must roll-off quickly above $\Omega_a$

$$\int_0^\infty \ln|S(j\omega)| \; d\omega = \pi p \quad \overset{\text{roughly}}{\Longrightarrow} \quad \Omega_S \leq \frac{\Omega_a \ln M - \pi p}{M}$$

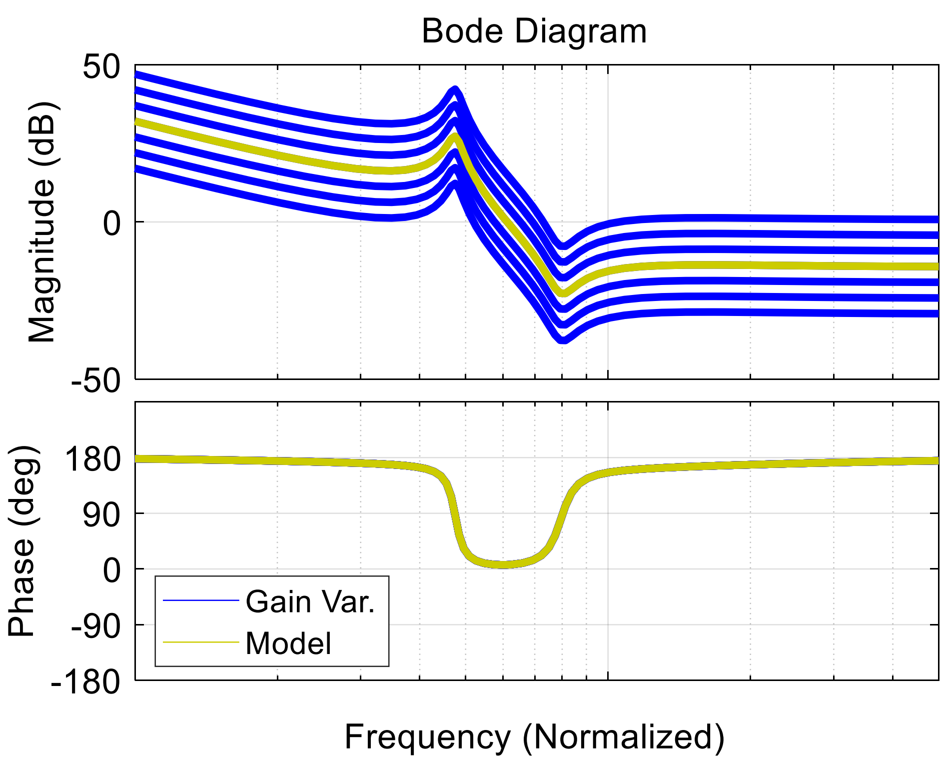
***Performance is constrained by the Bode integral and robustness requirements.***

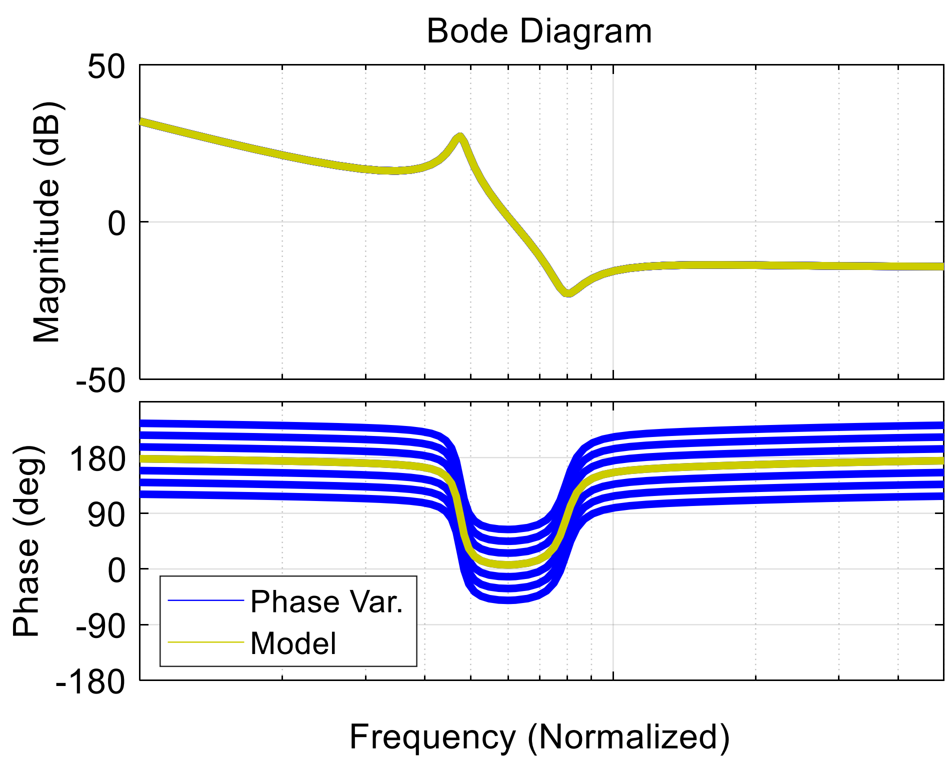


# Stability Margins: Safety Factors for Control

**Classical Margins:** Largest gain/phase variations that can be tolerated before closed-loop instability occurs.

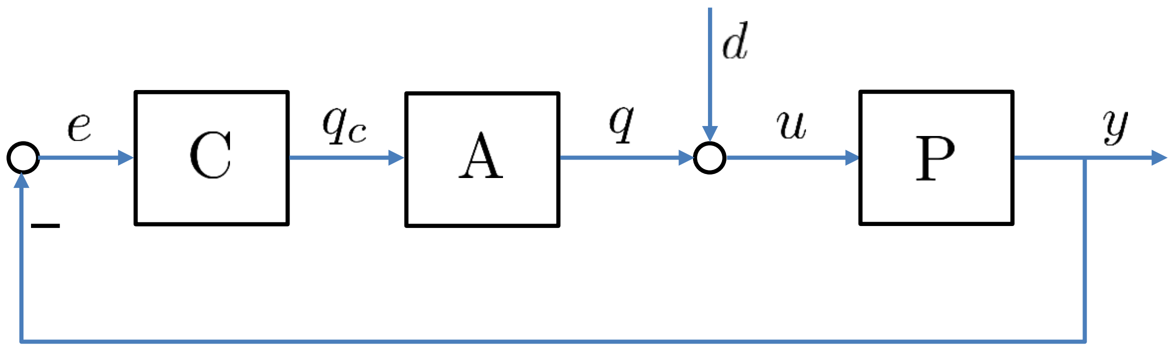
- Gain: $\alpha P$ where $\alpha$ varies from its nominal $\alpha_{nom}=1$
- Phase: $e^{j\theta}P$ where $\theta$ varies from its nominal $\theta_{nom}=0$



# Uncertainty Modeling

Consider SISO feedback system:



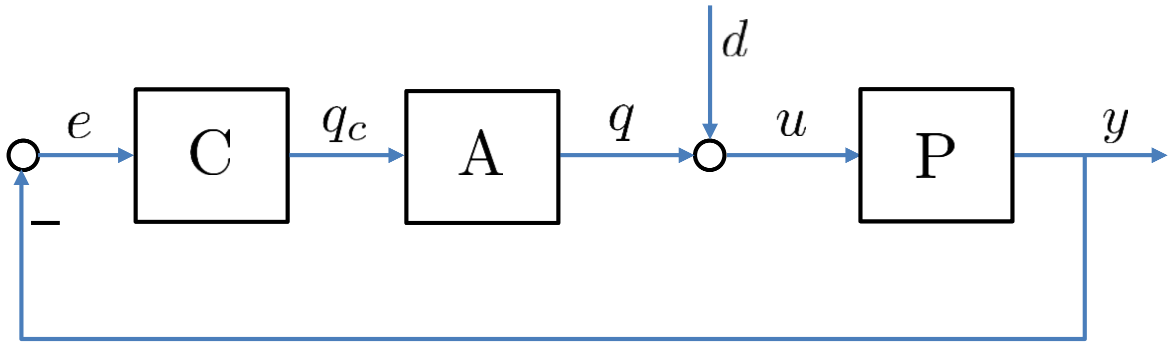

- Unstable plant with uncertain pole and input gain: $P(s) = \frac{b}{s-a}$ where $a \in [0.8, 1.1]$ and $b \in [1.7, 2.6]$

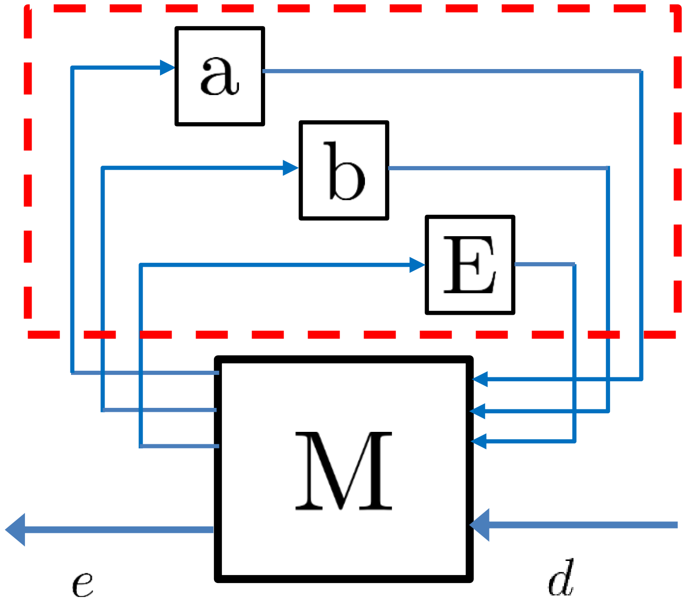
- First-order actuator with additive dynamic uncertainty $A(s) = A_0(s) + E(s)$ where $A_0(s) = \frac{10}{s+10}$ & $|E(\omega)| \leq 0.1,\ E$ stable

- Proportional-Integral control $C(s) = \frac{3s+4.5}{s}$

# Uncertainty Modeling

Separate known from the uncertain

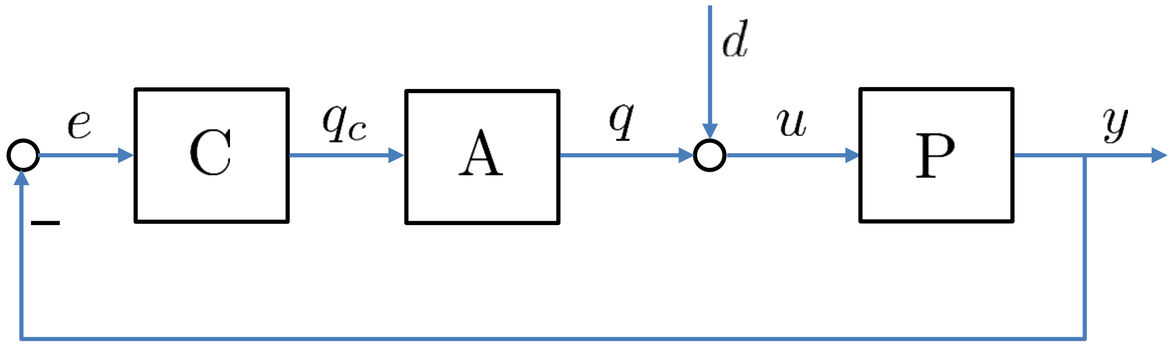


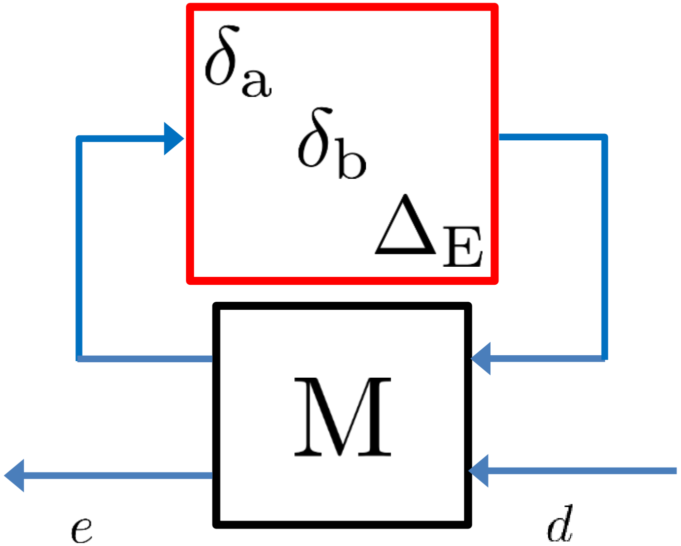




**Uncertainty is typically very structured**

# Uncertainty Modeling

Re-center and re-scale to normalize the uncertainties




Uncertainty set is structured:

$$\boldsymbol{\Delta} := \{\,\mathrm{diag}(\delta_a, \delta_b, \Delta_{\mathrm{E}}) \;:\; \delta_a, \delta_b \in \mathbb{R} \text{ and } \Delta_E \text{ LTI, stable}\}$$

where:

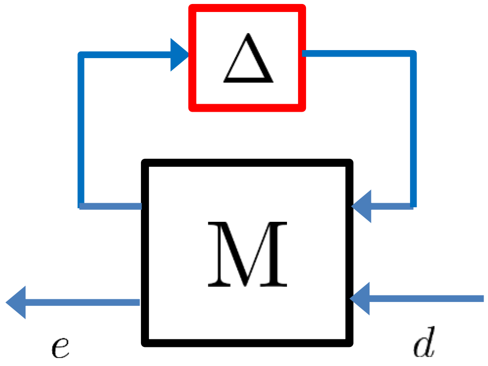
1. Δ=0 gives nominal behavior
2. Range of modeled uncertainty is

$$\|\Delta\|_\infty := \sup_\omega \bar{\sigma}(\Delta) \leq 1$$

# Robustness Metrics

**Stability Margin:** $\kappa_m := \inf_{\Delta \in \boldsymbol{\Delta}} \|\Delta\|_\infty$

$\text{s.t. } \Delta \text{ causes instability}$




**Worst-case Gain:** $\sup_{\substack{\Delta \in \boldsymbol{\Delta}, \\ \|\Delta\|_\infty \leq 1}} \| T_{d \to e}(M, \Delta) \|_\infty$

**Comments:**

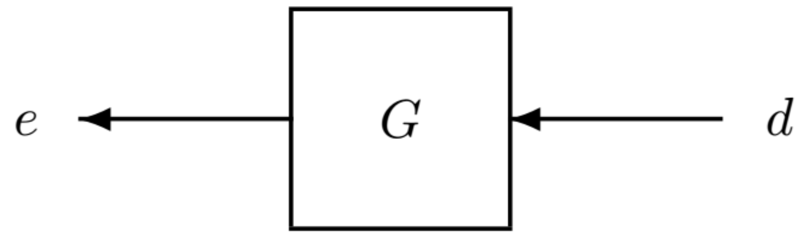
- System is robustly stable if and only if $\kappa_m > 1$.
- Both metrics can be converted to a (freq. domain) $\mu$ test.
- Algorithms compute bounds that provide guarantees on performance and bad instances of uncertainties.
- IQCs extend the framework to include nonlinearities.

# (Nominal) Finite Horizon Analysis

## Nominal LTV System

$$\begin{bmatrix} \dot{x}(t) \\ e(t) \end{bmatrix} = \begin{bmatrix} A(t) & B(t) \\ C(t) & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ d(t) \end{bmatrix}$$

$$x(0) = 0$$




## Analysis Objective

Derive bound on $\|e(T)\|_2$ that holds for all disturbances $\|d\|_{2,[0,T]} \leq 1$ on the horizon $[0, T]$.

# Nominal Analysis with Dissipation Inequalities

## Theorem [1,2]

If there exists $P(\cdot) = P(\cdot)^T$ such that
(i) $P(T) = C(T)^T C(T)$, and
(ii) $V(x,t) := x^T P(t) x$ satisfies

$$\frac{d}{dt} V(x,t) - \gamma^2 d(t)^T d(t) \leq 0 \ \ \forall t \in [0,T]$$

then $\|e(T)\|_2 \leq \gamma \|d\|_{2,[0,T]}$

## Proof

Integrate dissipation inequality from $t=0$ to $t=T$:

$$\underbrace{V(x(T),T)}_{=e(T)^T e(T)} - \underbrace{V(x(0),0)}_{=0} - \gamma^2 \int_0^T d(t)^T d(t) dt \leq 0$$

[1] Moore, Finite Horizon Robustness Analysis using IQCs, MS Thesis, Berkeley, 2015.
[2] Willems, Dissipative Dynamical Systems: Parts i and ii, 1972.

# Nominal Analysis with Dissipation Inequalities

**Theorem [1,2]**

If there exists $P(\cdot) = P(\cdot)^T$ such that
(i) $P(T) = C(T)^T C(T)$, and
(ii) $V(x,t) := x^T P(t) x$ satisfies

$$\frac{d}{dt} V(x,t) - \gamma^2 d(t)^T d(t) \leq 0 \;\; \forall t \in [0,T]$$

then $\|e(T)\|_2 \leq \gamma \|d\|_{2,[0,T]}$

Dissipation inequality can be recast as a differential LMI:

$$\begin{bmatrix} \dot{P} + A^T P + PA & PB \\ B^T P & -\gamma^2 I \end{bmatrix} \preceq 0 \;\; \forall t \in [0,T]$$

[1] Moore, Finite Horizon Robustness Analysis using IQCs, MS Thesis, Berkeley, 2015.
[2] Willems, Dissipative Dynamical Systems: Parts i and ii, 1972.

# Nominal Analysis with Dissipation Inequalities

## Theorem [1,2]

If there exists $P(\cdot) = P(\cdot)^T$ such that
$(i)$ $P(T) = C(T)^T C(T)$, and
$(ii)$ $V(x,t) := x^T P(t) x$ satisfies

$$\frac{d}{dt} V(x,t) - \gamma^2 d(t)^T d(t) \le 0 \;\; \forall t \in [0,T]$$

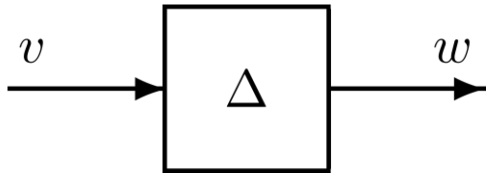
then $\|e(T)\|_2 \le \gamma \|d\|_{2,[0,T]}$

## Comments

- The dissipation inequality is equivalent to Riccati conditions [3] but enables extensions to robustness analysis.
- **Numerically reliable algorithm to construct worst-case disturbance [4].**

[1] Moore, Finite Horizon Robustness Analysis using IQCs, MS Thesis, Berkeley, 2015.
[2] Willems, Dissipative Dynamical Systems: Parts i and ii, 1972.
[3] Green & Limebeer, Linear Robust Control, 1995.
**[4] Iannelli, Seiler, Marcos, "Construction of worst-case disturbances for LTV systems…", 2019.**

# Example: Non-parametric (Dynamic) Uncertainty




$\Delta$ is stable, LTI with
$\|\Delta\|_\infty := \sup_\omega |\Delta(\omega)| \leq 1$

$|\hat{w}(\omega)| \leq |\hat{v}(\omega)| \, \forall \omega$ ⇒ $\int_{-\infty}^{\infty} X(\omega) \left[ |\hat{v}(\omega)|^2 - |\hat{w}(\omega)|^2 \right] \, d\omega \geq 0$
for any $X(\omega) \geq 0$

Frequency-Domain IQC ⇔

$$\int_{-\infty}^{\infty} \begin{bmatrix} \hat{v}(\omega) \\ \hat{w}(\omega) \end{bmatrix}^* \begin{bmatrix} X(\omega) & 0 \\ 0 & -X(\omega) \end{bmatrix} \begin{bmatrix} \hat{v}(\omega) \\ \hat{w}(\omega) \end{bmatrix} d\omega \geq 0$$

for any $X(\omega) \geq 0$

Spectral Factorization ⇔

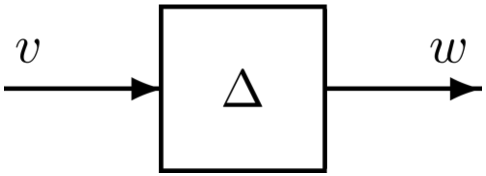
$X(\omega) = D(\omega)^* D(\omega)$

$$\int_{-\infty}^{\infty} \begin{bmatrix} \hat{v}(\omega) \\ \hat{w}(\omega) \end{bmatrix}^* \Psi(\omega)^* M \Psi(\omega) \begin{bmatrix} \hat{v}(\omega) \\ \hat{w}(\omega) \end{bmatrix} d\omega \geq 0$$

where

$\Psi := \text{diag}(D, D)$ and $M := \text{diag}(1, -1)$

# Example: Non-parametric (Dynamic) Uncertainty




$\Delta$ is stable, LTI with
$\|\Delta\|_\infty := \sup_\omega |\Delta(\omega)| \leq 1$

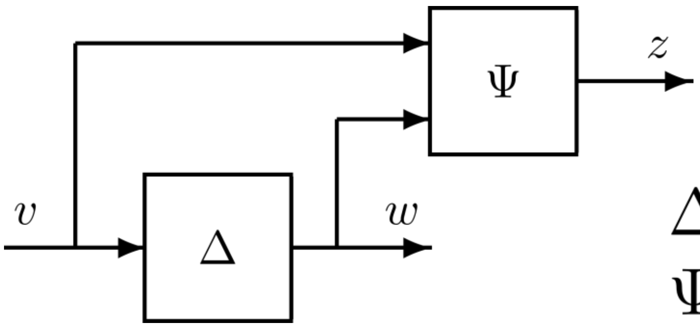
$|\hat{w}(\omega)| \leq |\hat{v}(\omega)| \, \forall \omega$ ⇒ $\int_{-\infty}^{\infty} X(\omega) \left[ |\hat{v}(\omega)|^2 - |\hat{w}(\omega)|^2 \right] \, d\omega \geq 0$
for any $X(\omega) \geq 0$

Frequency-Domain IQC ⇔

$$\int_{-\infty}^{\infty} \begin{bmatrix} \hat{v}(\omega) \\ \hat{w}(\omega) \end{bmatrix}^* \begin{bmatrix} X(\omega) & 0 \\ 0 & -X(\omega) \end{bmatrix} \begin{bmatrix} \hat{v}(\omega) \\ \hat{w}(\omega) \end{bmatrix} d\omega \geq 0$$

for any $X(\omega) \geq 0$

⇓ (Invoke Causality)




$\int_0^T z(t)^T M z(t) \, dt \geq 0$

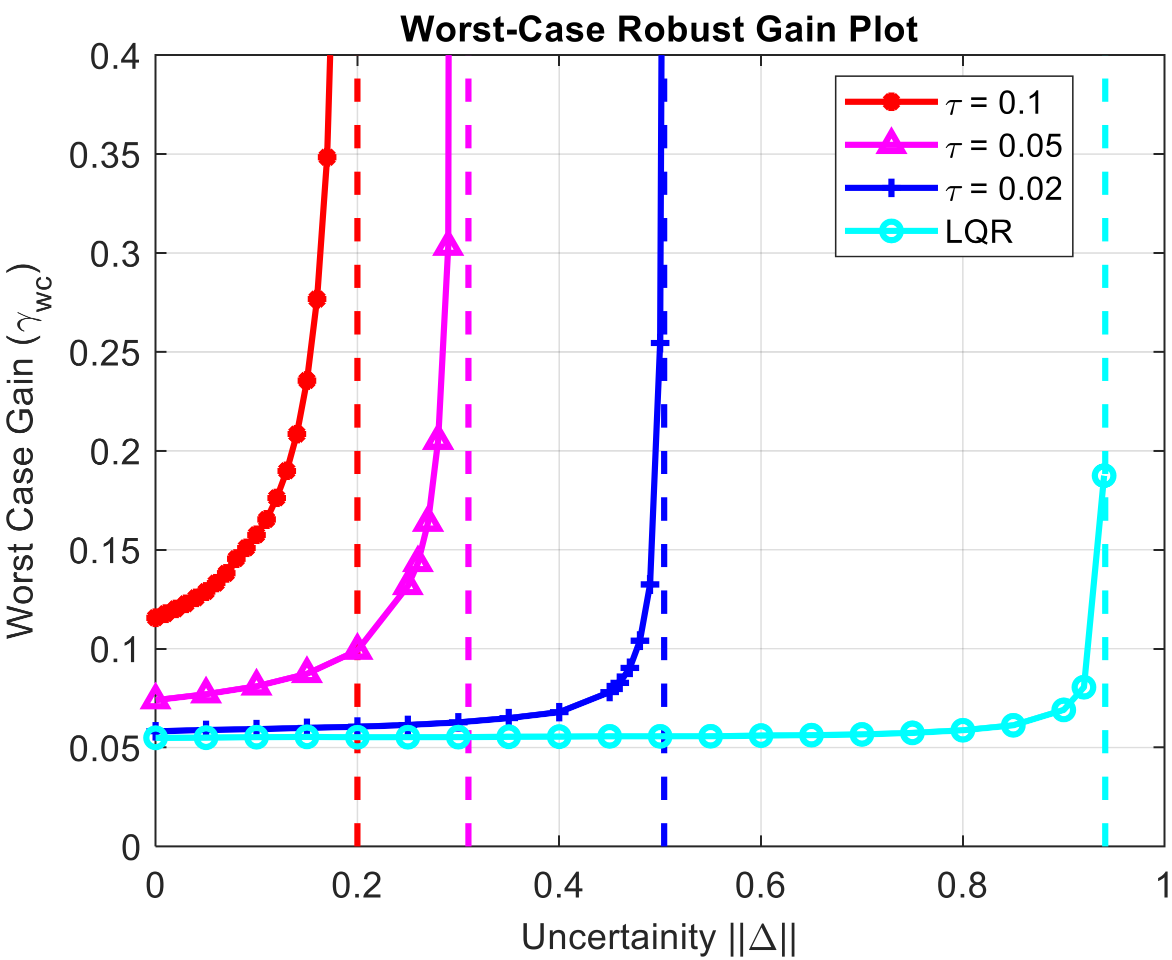
$\Delta$ satisfies the IQC on $[0, T]$ defined by $\Psi := \text{diag}(D, D)$ and $M := \text{diag}(1, -1)$

[1] Balakrishnan, Lyapunov Functionals in Complex μ Analysis, TAC, 2002.

# Closed-Loop Robust L2-to-Euclidean Gain

Two Controllers:

- Finite-Horizon LQR with state feedback
- Output Feedback using high pass filter $\tau s/(\tau s+1)$ to estimate angular rates

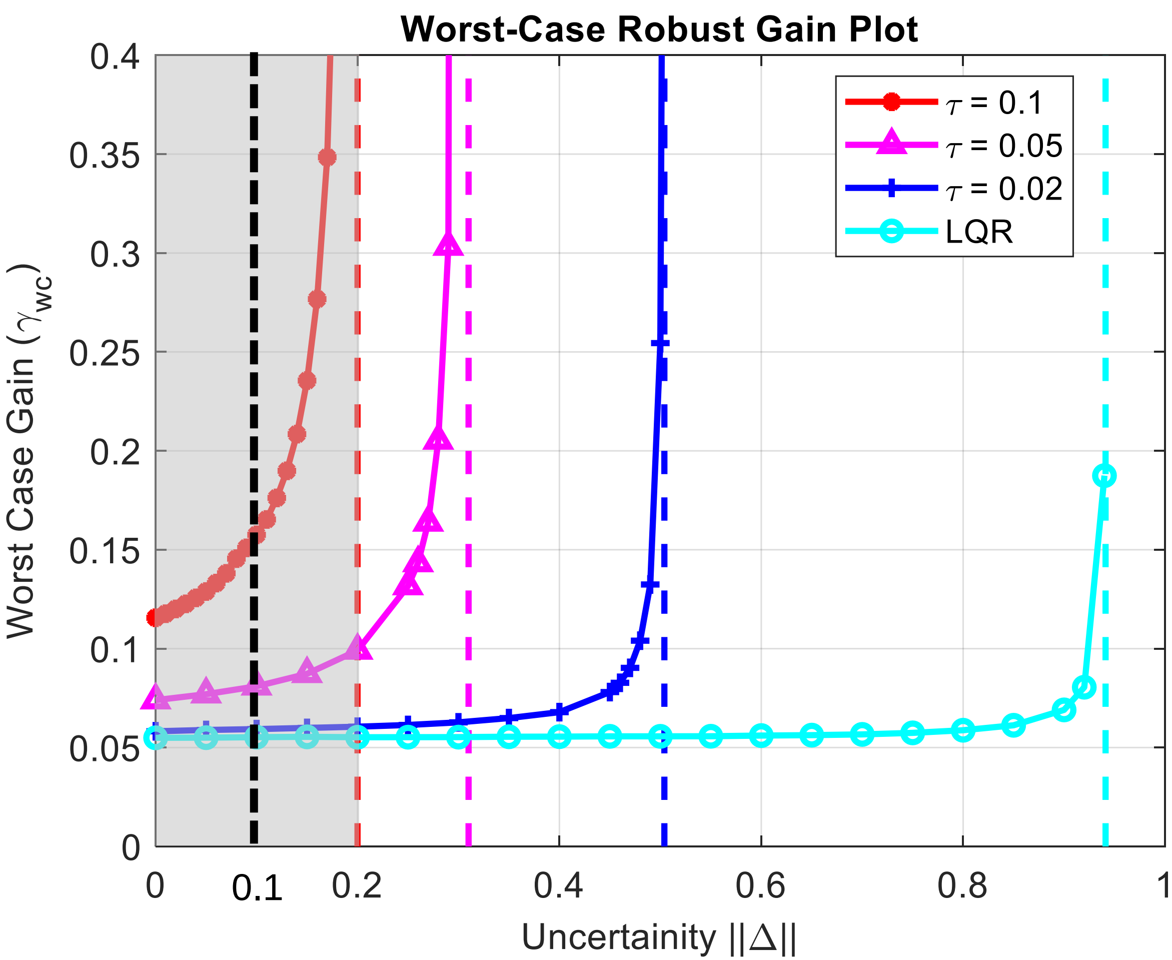



Finite horizon robustness is degraded by output feedback with rate estimates.

# Closed-Loop Robust L2-to-Euclidean Gain

Two Controllers:

- Finite-Horizon LQR with state feedback
- Output Feedback using high pass filter $\tau s/(\tau s+1)$ to estimate angular rates

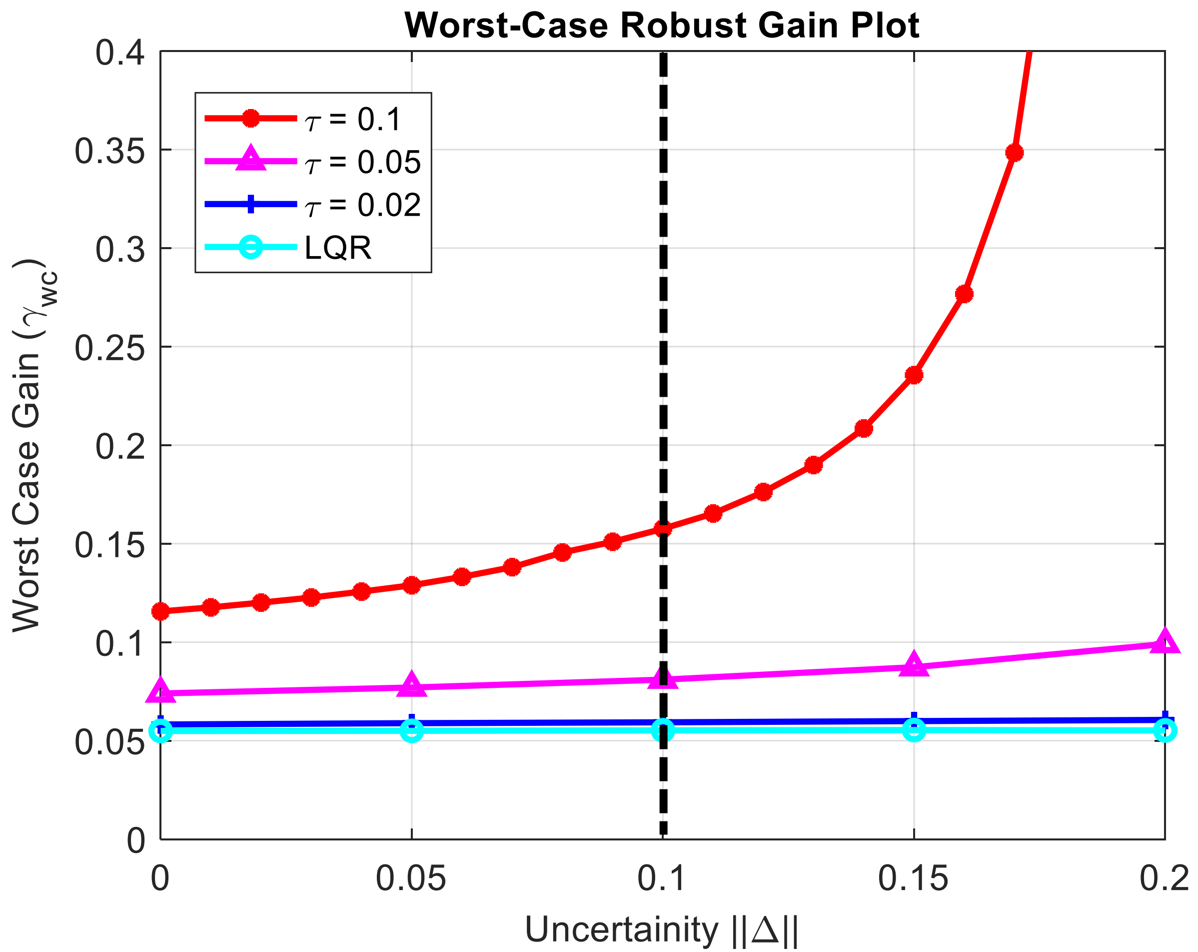



Finite horizon robustness is degraded by output feedback with rate estimates.

# Closed-Loop Robust L2-to-Euclidean Gain

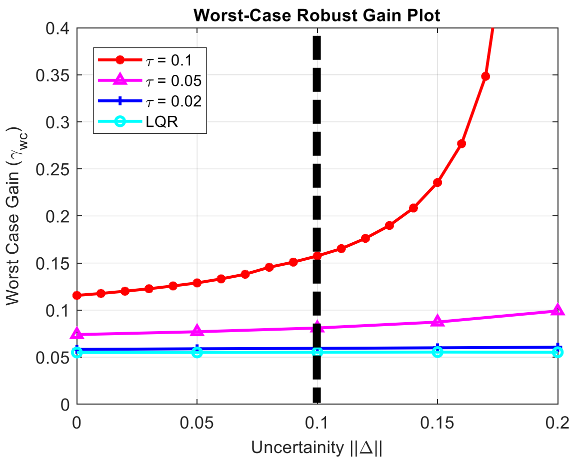
Two Controllers:

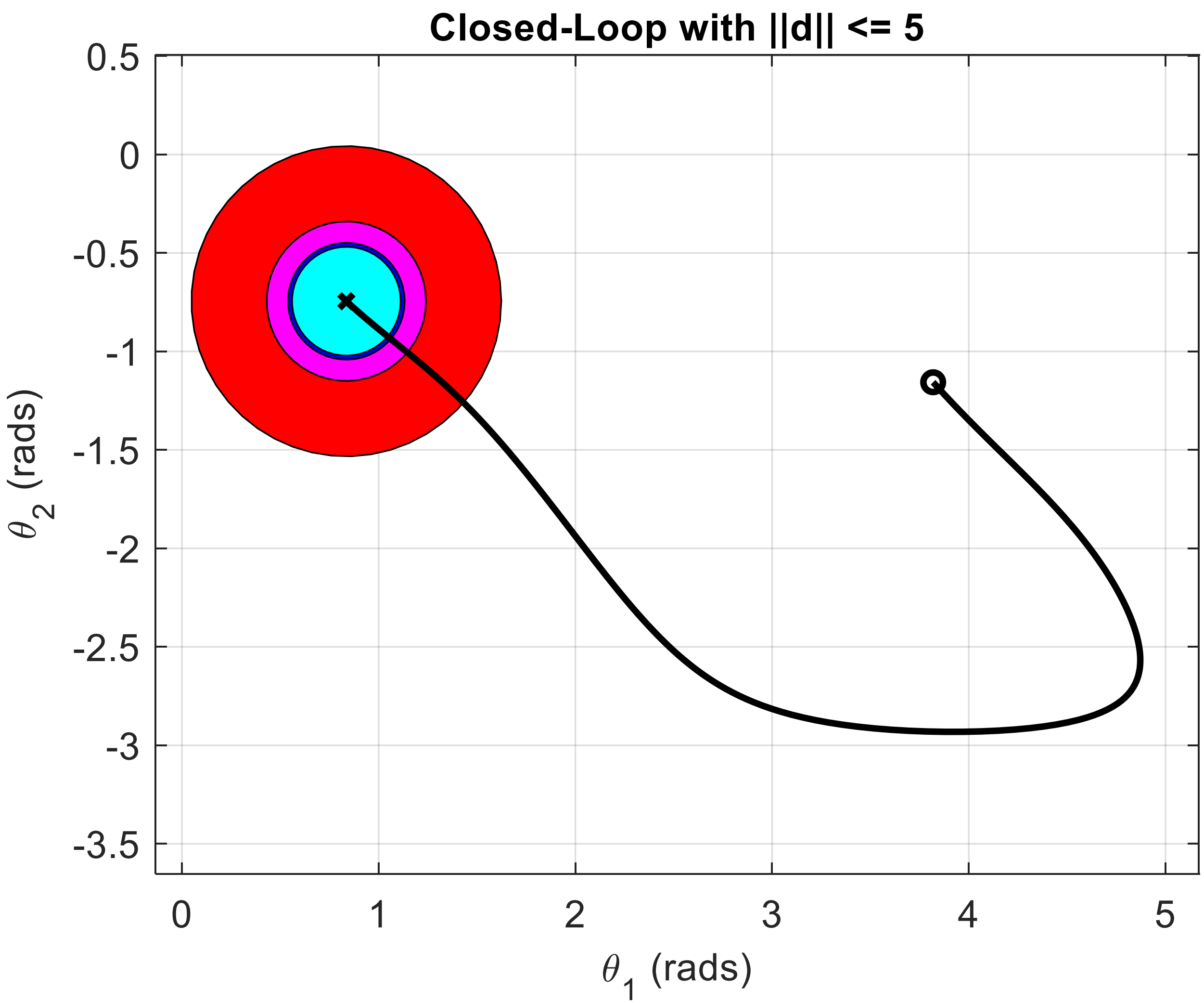
- Finite-Horizon LQR with state feedback
- Output Feedback using high pass filter *τs/(τs+1)* to estimate angular rates




Finite horizon robustness is degraded by output feedback with rate estimates.

# Impact of Using High Pass Rate Estimator



# Partial Dictionary of IQCs [1]

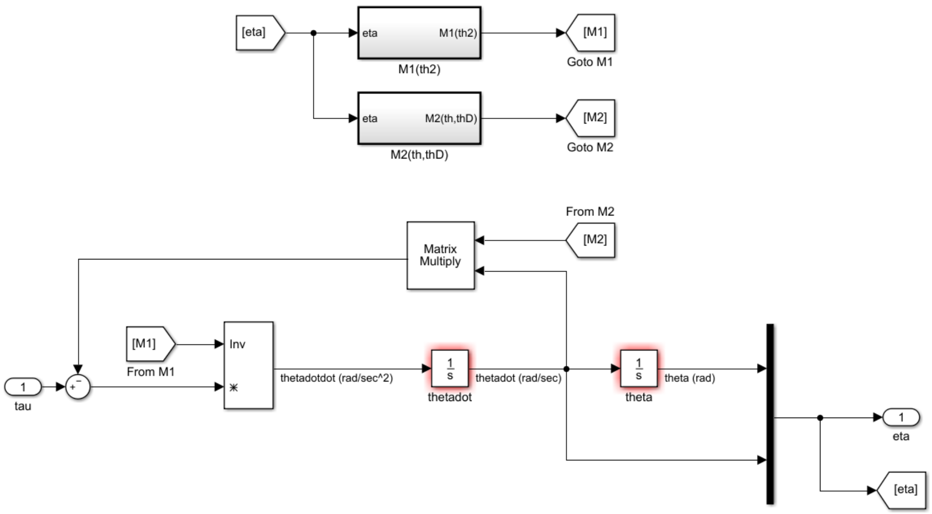
| Uncertainty | IQC Multiplier |
|---|---|
| 1. Passive | $\begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$ |
| 2. Norm-bounded LTI | $\begin{bmatrix} X(j\omega) & 0 \\ 0 & -X(j\omega) \end{bmatrix}$ where $X(j\omega) \geq 0$ |
| 3. Constant Real Parameter | $\begin{bmatrix} X(j\omega) & Y(j\omega) \\ Y(j\omega)^* & -X(j\omega) \end{bmatrix}$ where $X(j\omega) \geq 0$ and $Y(j\omega) = -Y(j\omega)^*$. |
| 4. Varying Real Parameter | $\begin{bmatrix} X & Y \\ Y^T & -X \end{bmatrix}$ where $X \geq 0$ and $Y = -Y^T$. |
| 5. Unit Saturation | $\begin{bmatrix} 0 & 1 + H(j\omega) \\ 1 + H(j\omega)^* & -2(1 + ReH(j\omega)) \end{bmatrix}$ where $\\|h\\|_1 \leq 1$. |

[1] Megretski & Rantzer, System analysis via IQCs, TAC, 1997. [IQCs derived based on much prior literature]

# LTV Toolchain

```
% Matlab snapshot linearizations
% along nominal trajectory
io(1) = linio('TwoLinkRobotOL/Input Torque',1,'input');
io(2) = linio('TwoLinkRobotOL/Two Link Robot Arm',1,'output');
sys = linearize('TwoLinkRobotOL',io,Tgrid);

% Construction of LTV Model
G = tvss(sys,Tgrid);
```




Simulink Model of Robotic Arm

# Summary: Recovering Robustness in RL

- Robustness issues can arise in output-feedback controllers trained by RL [2]
  - Linear Quadratic Gaussian (LQG) Control can be solved via RL
  - A well-known counterexample by Doyle [1] demonstrates that LQG controllers can have arbitrarily small margins.

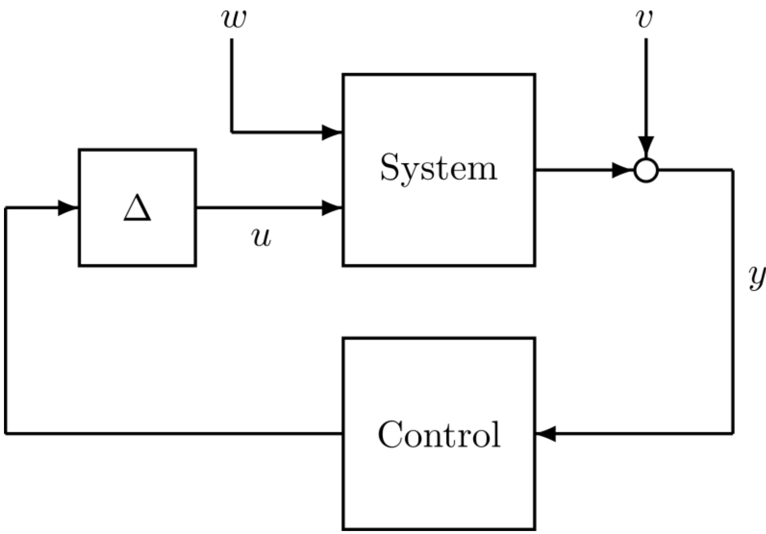
[1] J. Doyle. Guaranteed margins for LQG regulators, IEEE TAC, 1978.

[2] Venkataraman & Seiler, Recovering Robustness in Model-Free Reinforcement Learning, '18 arXiv and '19 ACC submission.

# Summary: Recovering Robustness in RL

- Robustness issues can arise in output-feedback controllers trained by RL [2]
  - Linear Quadratic Gaussian (LQG) Control can be solved via RL
  - A well-known counterexample by Doyle [1] demonstrates that LQG controllers can have arbitrarily small margins.

- Robustness can be recovered by introducing (synthetic) input perturbations during the RL training [2].





[1] J. Doyle. Guaranteed margins for LQG regulators, IEEE TAC, 1978.

[2] Venkataraman & Seiler, Recovering Robustness in Model-Free Reinforcement Learning, '18 arXiv and '19 ACC submission.