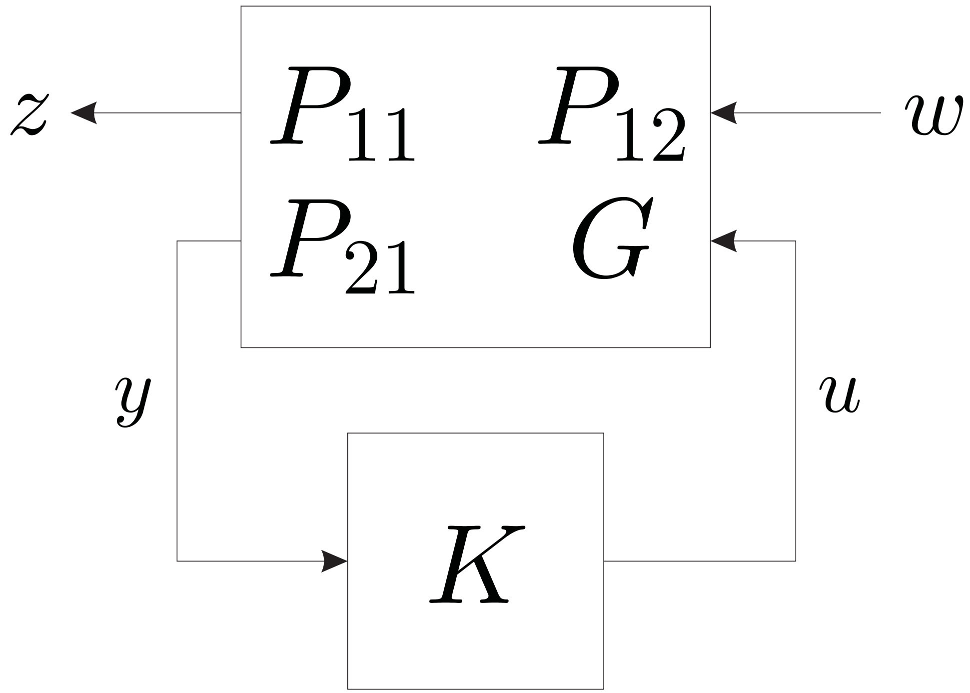
When is a Linear Controller Optimal?

Michael Rotkowitz

Research School of Information Sciences and Engineering
The Australian National University

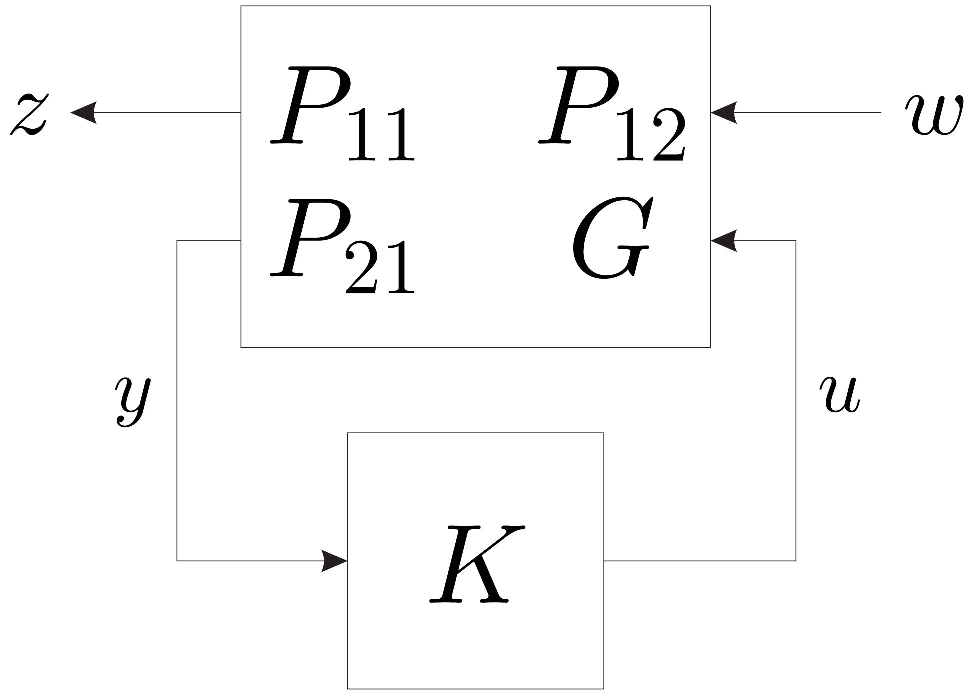
LinSys 2007
Canberra, Australia
28 February 2007

General Formulation



minimize $\|f_l(P,K)\|$

General Formulation



$$\begin{array}{rl} \text{minimize} & \|f_l(P,K)\| \\ \text{subject to} & K \in S \end{array}$$

Uniform Optimal (Centralized) Control

$$\text{minimize} \qquad \|f_l(P,K)\|_\infty$$

- where

$$\|G\|_\infty = \sup_{w\neq 0} \frac{\|Gw\|_2}{\|w\|_2}$$

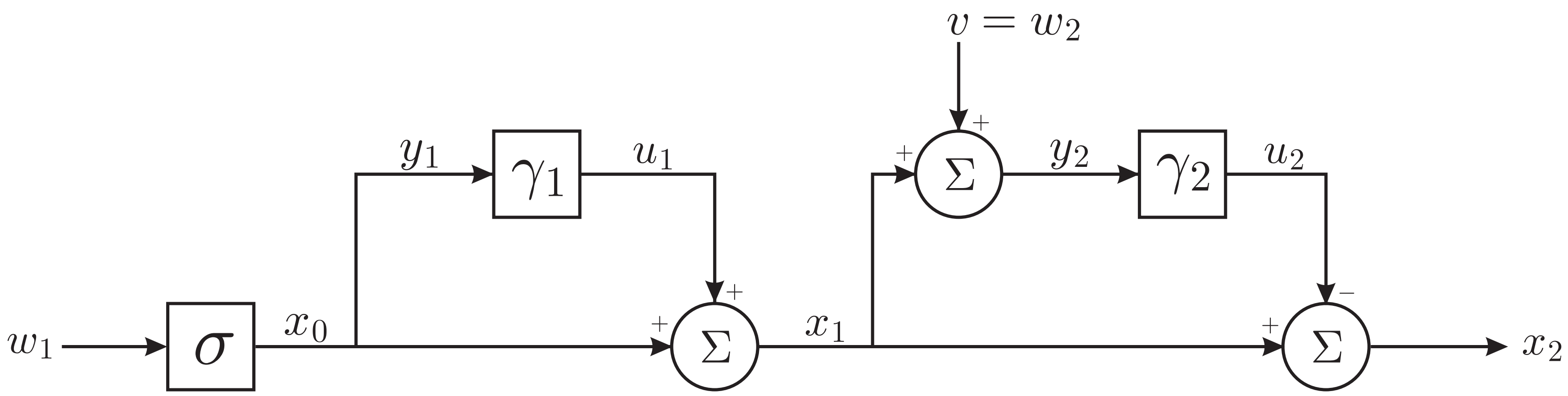
- When plant P is linear, optimal controller K is indeed linear
  - Feintuch and Francis, 1985
  - Khargonekar and Poolla, 1986

LQG

$$\text{minimize} \quad \mathbb{E}\left(h\left(z\right)\right) \quad \text{where } z = f_l\left(P, K\right)w$$

- "strictly classical information pattern" $\Rightarrow$ separation
  - $\gamma^* = \phi \circ F$
- Gaussian noise $\Rightarrow$ linear estimation
  - F^* linear
- quadratic cost $\Rightarrow$ linear control
  - ϕ^* linear
- Witsenhausen, 1971

Witsenhausen Counterexample (1968)



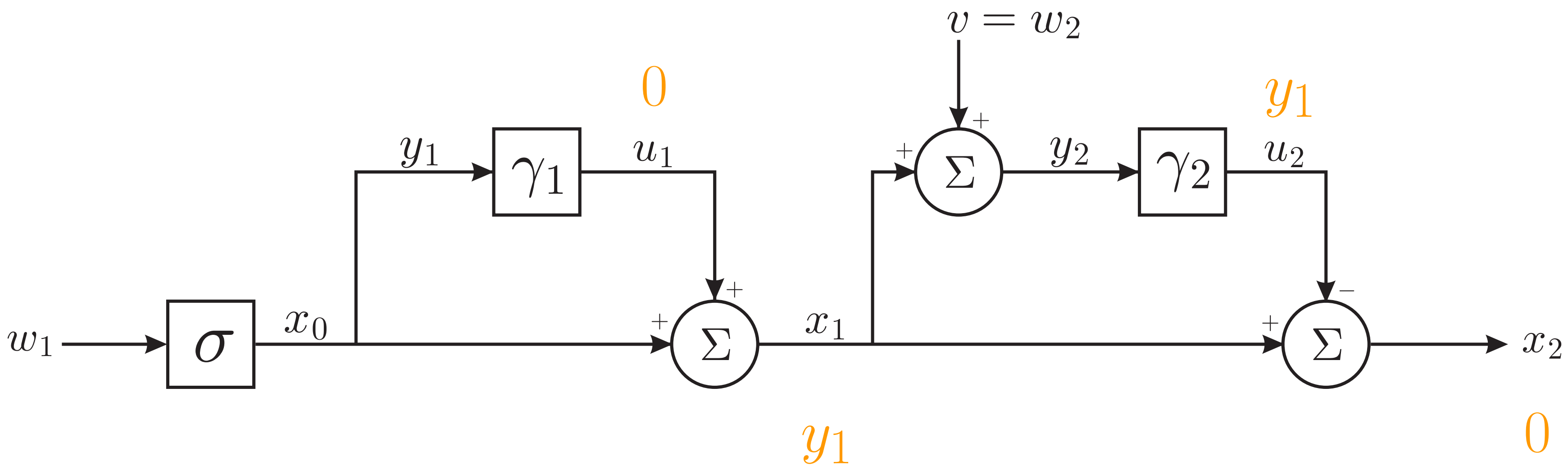
Objective was to seek γ_1, γ_2 to minimize

$$\mathbb{E}\left(k^2 u_1^2 + x_2^2\right)$$

where

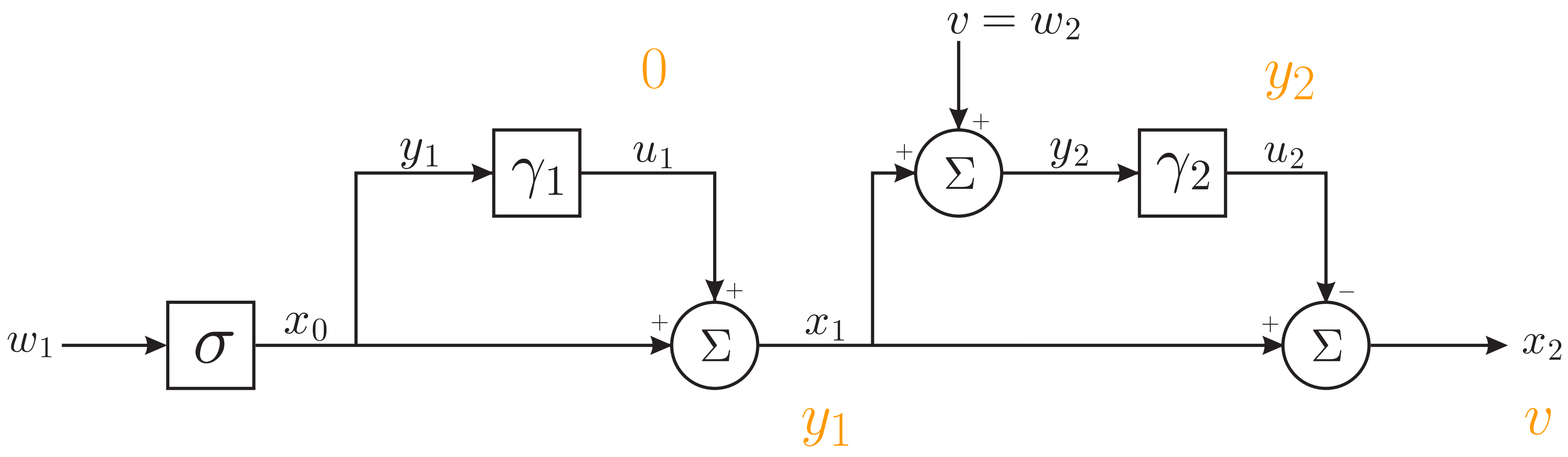
$$w_1 \sim \mathcal{N}(0,1) \qquad\qquad w_2 \sim \mathcal{N}(0,1)$$

W.C. - Full Information



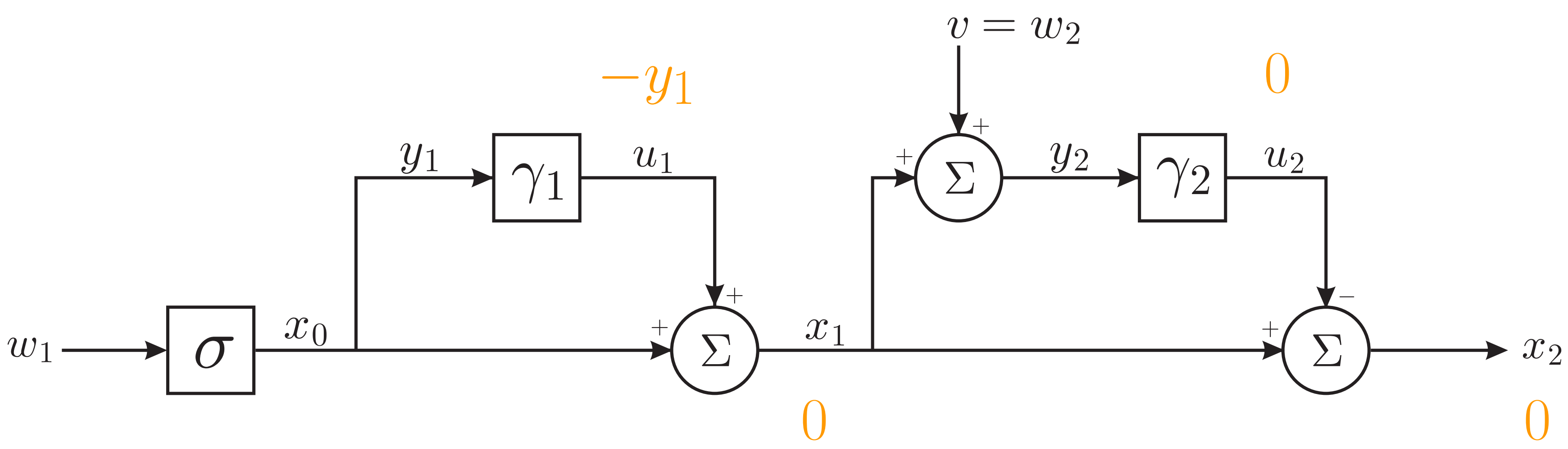
	γ_1	γ_2	$E(u_1^2)$	x_2	$E(x_2^2)$	J
(0) full info	0	y_1	0	0	0	0

W.C. - No Input Cost



	γ_1	γ_2	$E(u_1^2)$	x_2	$E(x_2^2)$	J
(0) full info	0	y_1	0	0	0	0
(1) no input cost	0	y_2	0	v	1	1

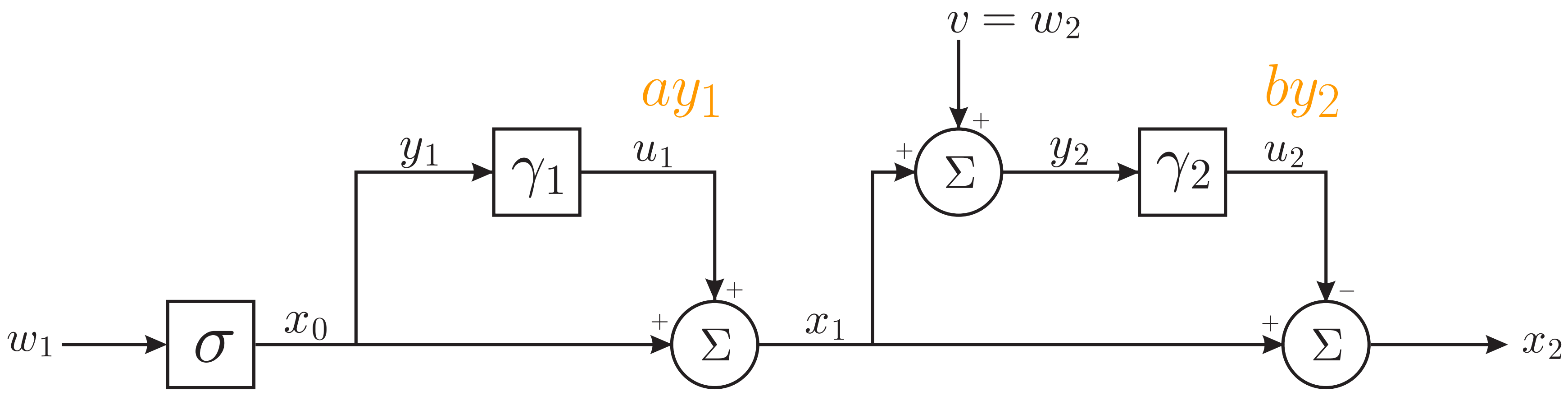
W.C. - No Output Cost



Let $k = 0.1$ and $\sigma = 10$.

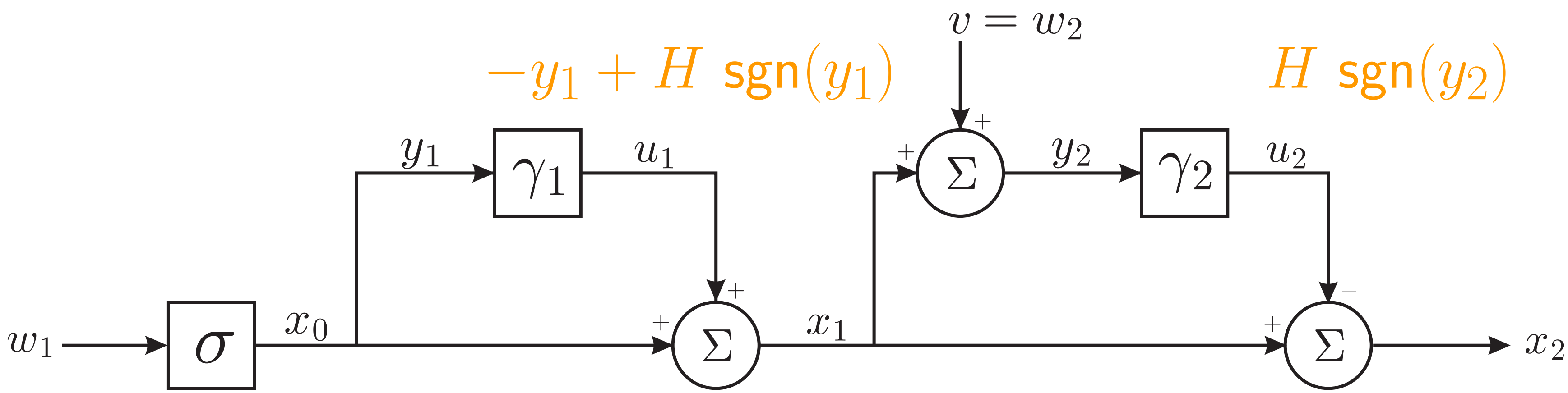
	γ_1	γ_2	$E(u_1^2)$	x_2	$E(x_2^2)$	J
(0) full info	0	y_1	0	0	0	0
(1) no input cost	0	y_2	0	v	1	1
(2) no output cost	$-y_1$	0	$\sigma^2 = 100$	0	0	$k^2\sigma^2 = 1$

W.C. - Best Linear



	γ_1	γ_2	$E(u_1^2)$	x_2	$E(x_2^2)$	J
(0) full info	0	y_1	0	0	0	0
(1) no input cost	0	y_2	0	v	1	1
(2) no output cost	$-y_1$	0	100	0	0	1
(3) best linear	$-0.01y_1$	$0.99y_2$	0.01	-	0.9899	0.99

W.C. - Signalling

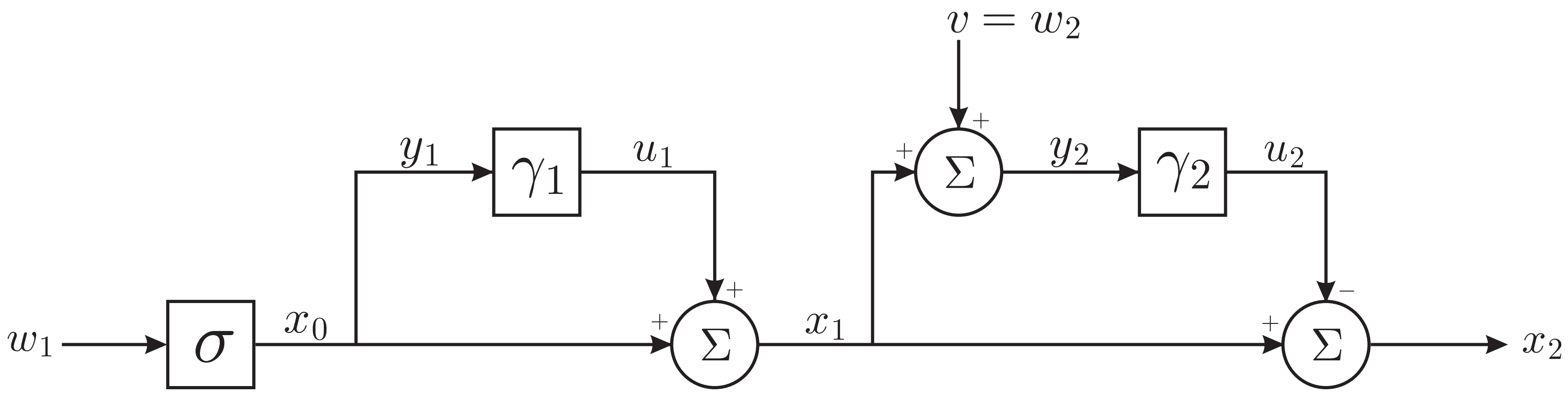


$$\begin{cases} +H, & \text{if } y_1 \geq 0 \\ -H, & \text{if } y_1 < 0 \end{cases}$$

$$\begin{cases} +2H, & \text{if } w_1 \geq 0,\ w_2 < -H \\ -2H, & \text{if } w_1 < 0,\ w_2 \geq +H \\ 0, & \text{otherwise} \end{cases}$$

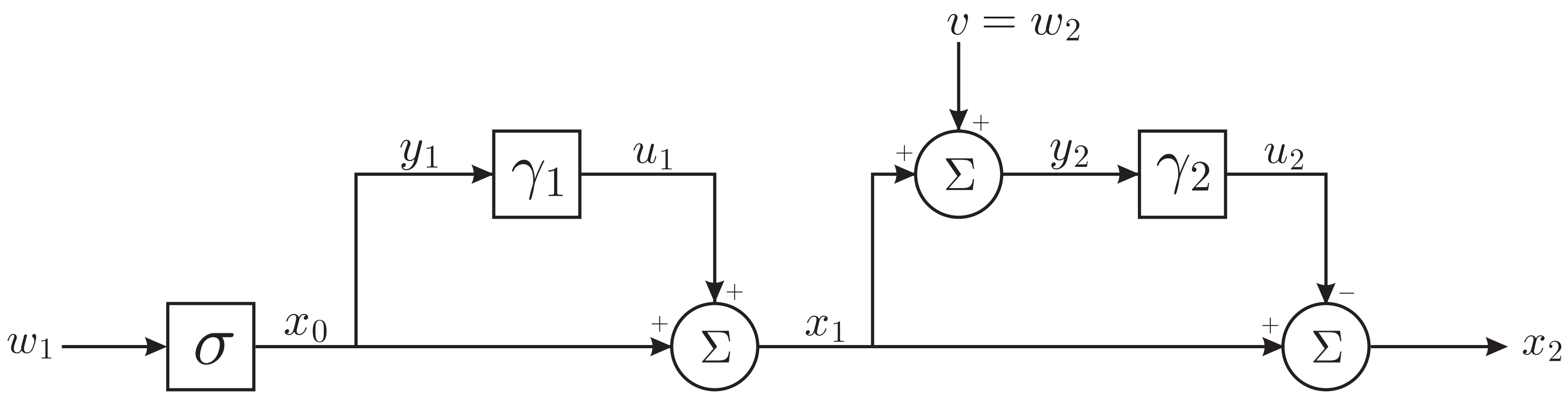
	γ_1	γ_2	$E(u_1^2)$	x_2	$E(x_2^2)$	J
(0) full info	0	y_1	0	0	0	0
(1) no input cost	0	y_2	0	v	1	1
(2) no output cost	$-y_1$	0	100	0	0	1
(3) best linear	ay_1	by_2	0.01	-	0.9899	0.99
(4) M&S	-	H sgn(y_2)	40	$\{0, \pm 2H\}$	$\backsim 0$	0.40

W.C. - Hierarchical Search



	γ_1	γ_2	$E(u_1^2)$	x_2	$E(x_2^2)$	J
(0) full info	0	y_1	0	0	0	0
(1) no input cost	0	y_2	0	v	1	1
(2) no output cost	$-y_1$	0	σ^2	0	0	1
(3) best linear	ay_1	by_2	0.01	-	0.9899	0.99
(4) M&S	-	$H\ \text{sgn}(y_2)$	40	$\{0, \pm 2H\}$	$\backsim 0$	0.40
(5) LLH	-	-	13.2	-	0.035	0.167

Witsenhausen Counterexample (1968)



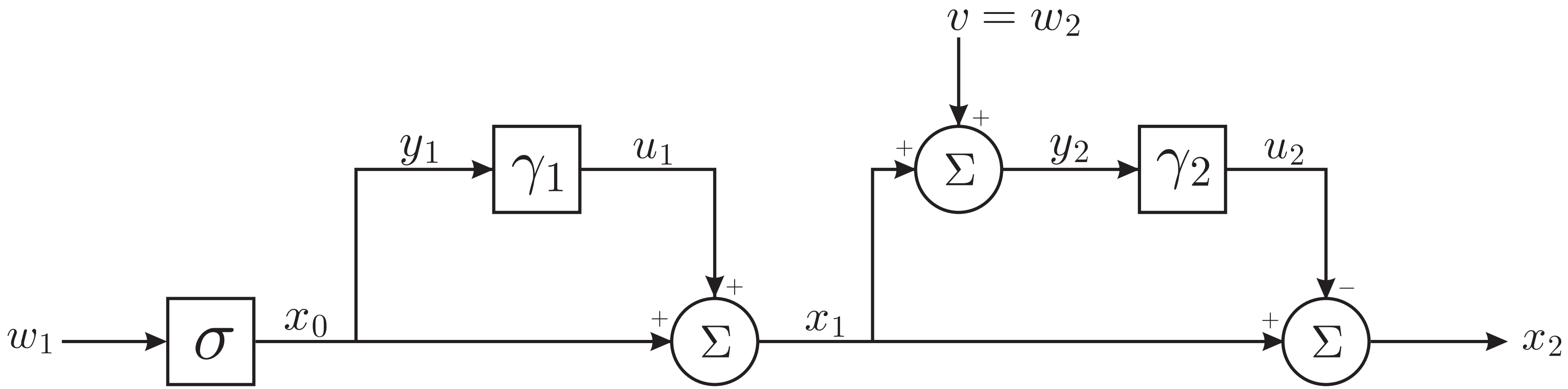
Objective was to seek γ_1, γ_2 to minimize

$$\mathbb{E}\left(k^2 u_1^2 + x_2^2\right)$$

where

$$w_1 \sim \mathcal{N}(0,1) \qquad\qquad w_2 \sim \mathcal{N}(0,1)$$

Witsenhausen Counterexample (1968)

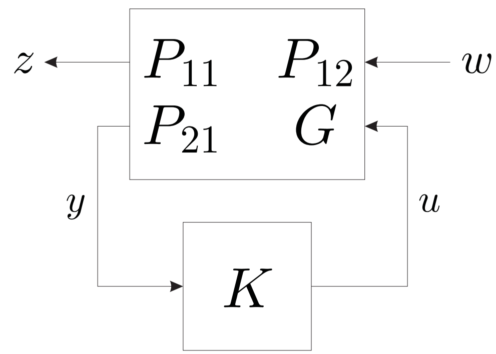


Letting

$$w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \qquad z = \begin{bmatrix} k\ u_1 \\ x_2 \end{bmatrix}$$

we then seek to minimize

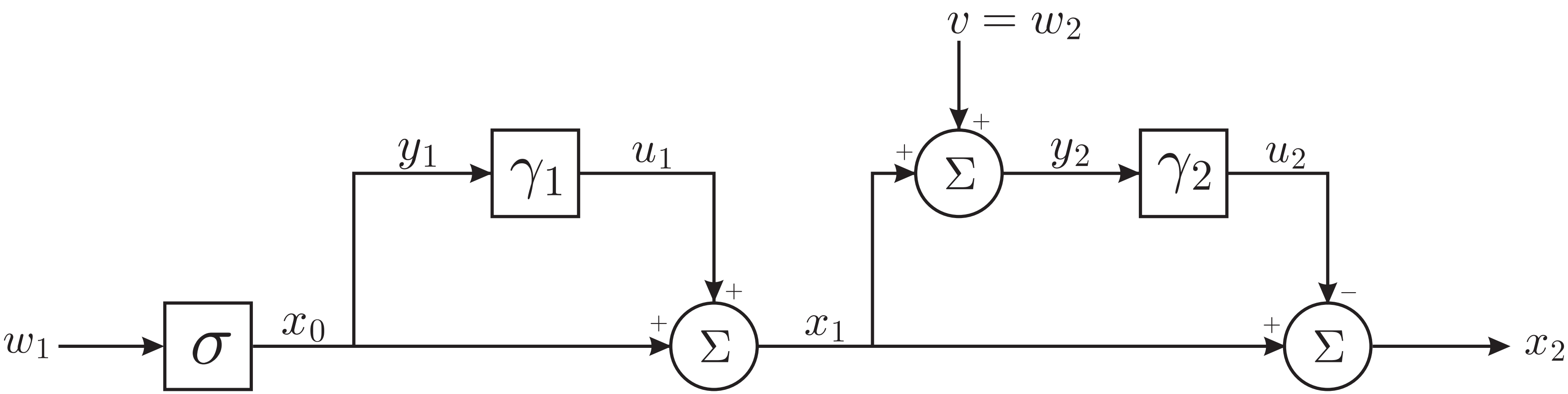
$$J(\gamma_1, \gamma_2) \ = \ \mathbb{E}\ \|z\|_2^2$$



where

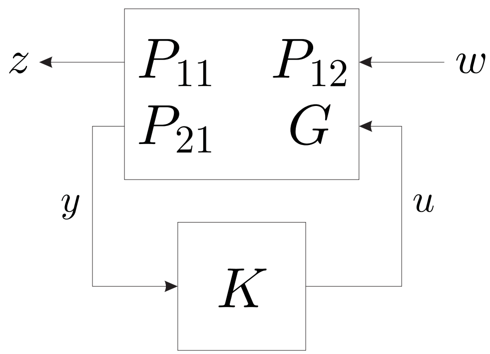
$$w \sim \mathcal{N}(0, I)$$

Witsenhausen Counterexample - Induced Norm



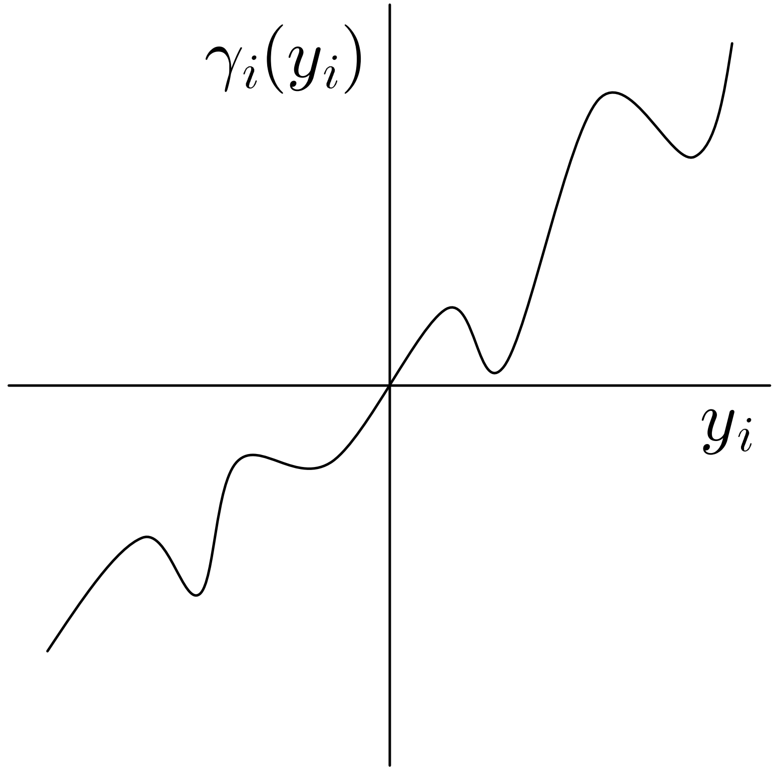
Instead consider the induced norm, and seek γ_1, γ_2 to minimize

$$J(\gamma_1, \gamma_2) \;=\; \sup_{w \neq 0} \frac{\|z\|_2}{\|w\|_2}$$



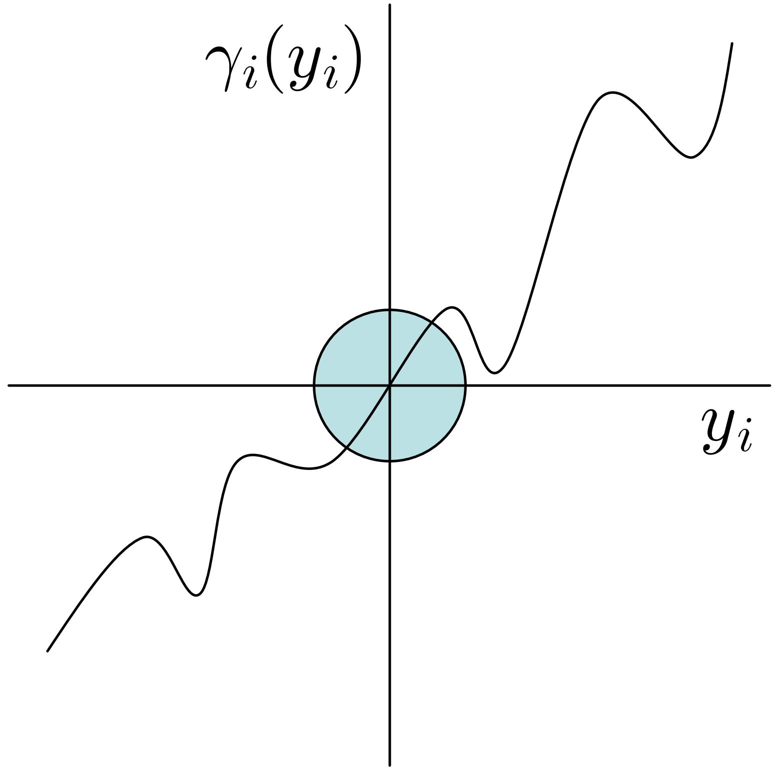
Induced Norms

$$\sup_{w \neq 0} \frac{\|z\|_2}{\|w\|_2}$$



Induced Norms

$$\sup_{w \neq 0} \frac{\|z\|_2}{\|w\|_2} \;\geq\; \sup_{\|w\| \leq \epsilon} \frac{\|z\|_2}{\|w\|_2}$$

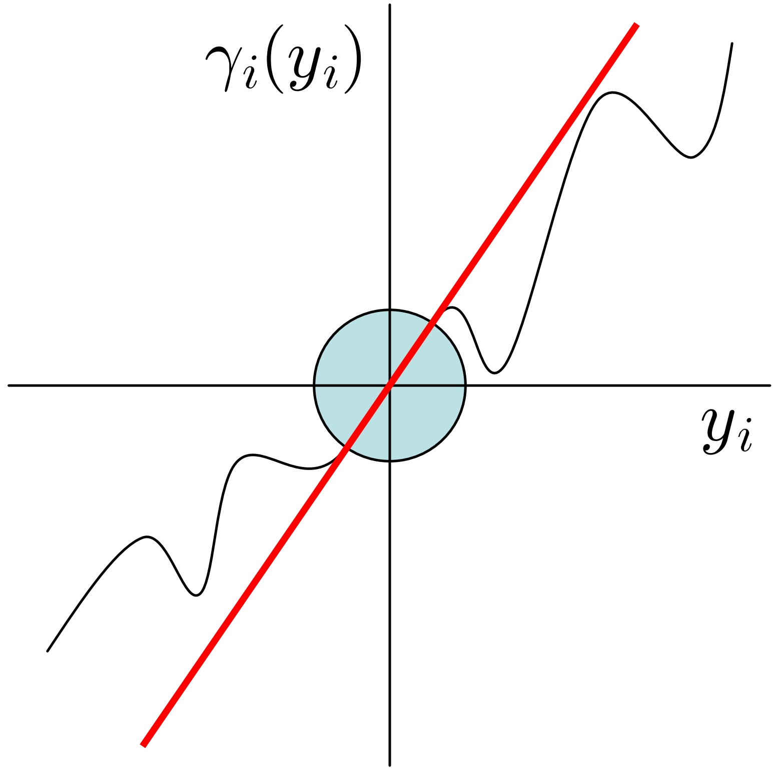


Induced Norms

$$\sup_{w \neq 0} \frac{\|z\|_2}{\|w\|_2} \geq \sup_{\|w\| \leq \epsilon} \frac{\|z\|_2}{\|w\|_2}$$

$$a = \left.\frac{d\gamma_1}{dy_1}\right|_{y_1=0} \qquad b = \left.\frac{d\gamma_2}{dy_2}\right|_{y_2=0}$$

$$\gamma_1 \approx ay_1 \qquad \gamma_2 \approx by_2$$



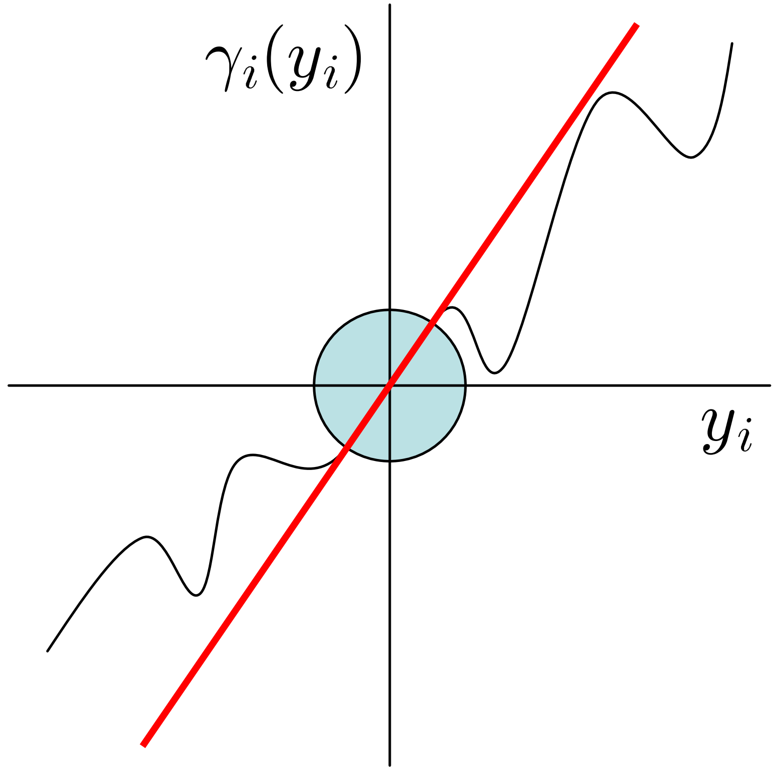
Induced Norms

$$\sup_{w \neq 0} \frac{\|z\|_2}{\|w\|_2} \geq \sup_{\|w\| \leq \epsilon} \frac{\|z\|_2}{\|w\|_2}$$

$$a = \left. \frac{d\gamma_1}{dy_1} \right|_{y_1=0} \qquad b = \left. \frac{d\gamma_2}{dy_2} \right|_{y_2=0}$$

$$\gamma_1 \approx a y_1 \qquad \gamma_2 \approx b y_2$$

$$J(\gamma_1, \gamma_2) = \sup_{w \neq 0} \frac{\|z\|_2}{\|w\|_2} \geq J_l(a, b)$$

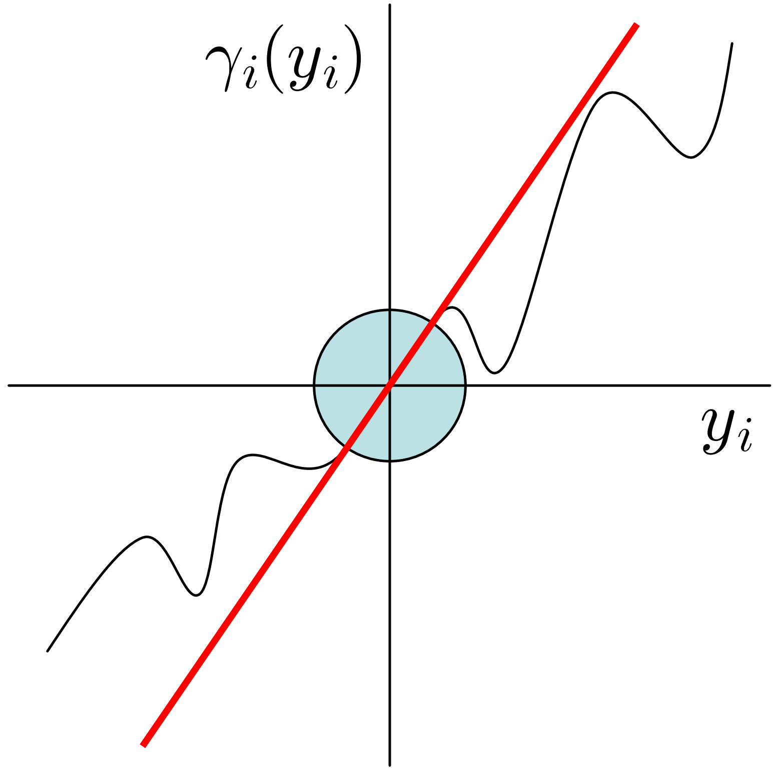


Induced Norms

$$\sup_{w \neq 0} \frac{\|z\|_2}{\|w\|_2} \geq \sup_{\|w\| \leq \epsilon} \frac{\|z\|_2}{\|w\|_2}$$

$$a = \left.\frac{d\gamma_1}{dy_1}\right|_{y_1=0} \qquad b = \left.\frac{d\gamma_2}{dy_2}\right|_{y_2=0}$$

$$\gamma_1 \approx ay_1 \qquad \gamma_2 \approx by_2$$



$$J(\gamma_1, \gamma_2) = \sup_{w \neq 0} \frac{\|z\|_2}{\|w\|_2} \geq J_l(a, b)$$

May as well fix

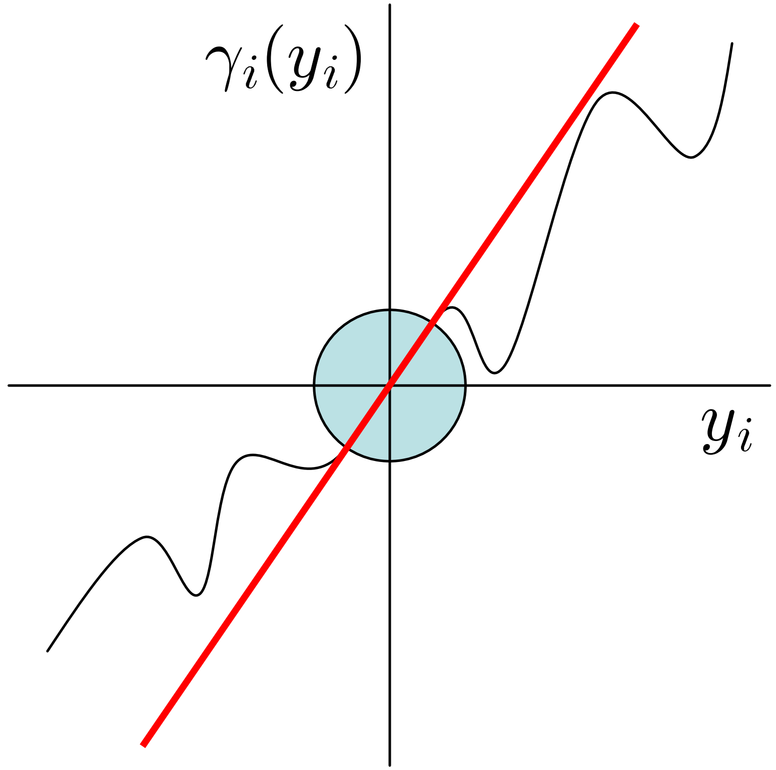
$$\gamma_1(y_1) = ay_1 \qquad \gamma_2(y_2) = by_2$$

Can't do any better!

Induced Norms

$$\gamma_1 \approx ay_1 \qquad \gamma_2 \approx by_2$$

$$J(\gamma_1, \gamma_2) = \sup_{w \neq 0} \frac{\|z\|_2}{\|w\|_2} \geq J_l(a, b)$$



May as well fix

$$\gamma_1(y_1) = ay_1 \qquad \gamma_2(y_2) = by_2$$

Can't do any better!

- Given any nonlinear controller that's differentiable in the origin, we have a linear controller that's at least as good.
- End of story? Not so fast.

Another Induced Norm

- L_1 / ℓ_1 $(\infty \to \infty)$ Norm

$$\sup_{w \neq 0} \frac{\|z\|_\infty}{\|w\|_\infty}$$

- (M. Dahleh and J. Shamma, 1992) Nonlinear controllers which are differentiable in the origin cannot do better than linear controllers.

- But............

A. Stoorvogel, 1995

$$
\begin{aligned}
z &= \begin{bmatrix} 0 & 3 & 0 \\ -1.5 & 1.5 & -3 \\ -3 & 0 & 0 \\ 0 & 0 & -3 \end{bmatrix} w + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} u \\
y &= w
\end{aligned}
$$

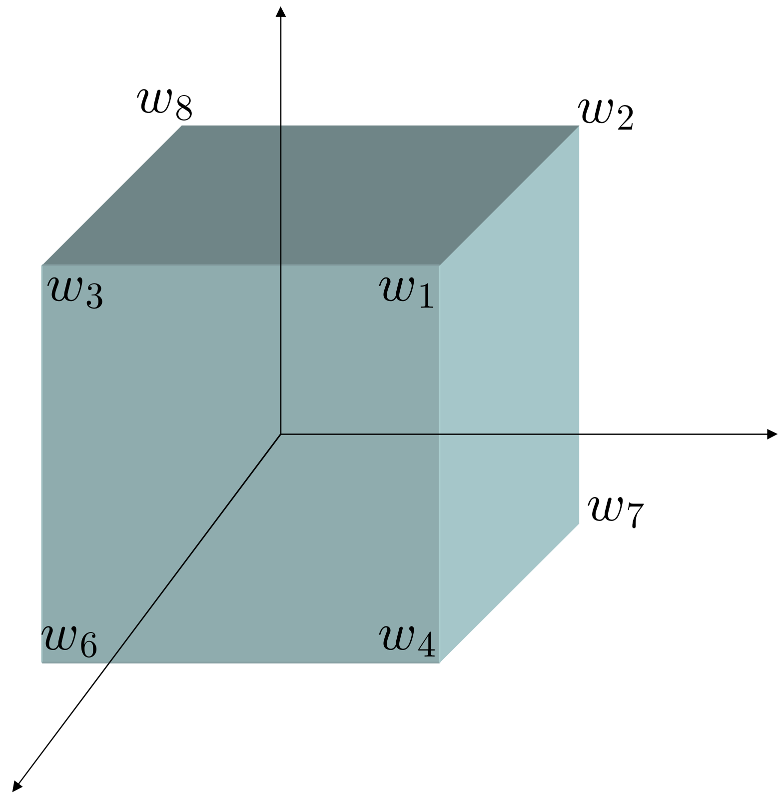
$$
u = Ky
$$

Find K to minimize

$$
J(K) = \sup_{w \neq 0} \frac{\|z\|_\infty}{\|w\|_\infty}
$$

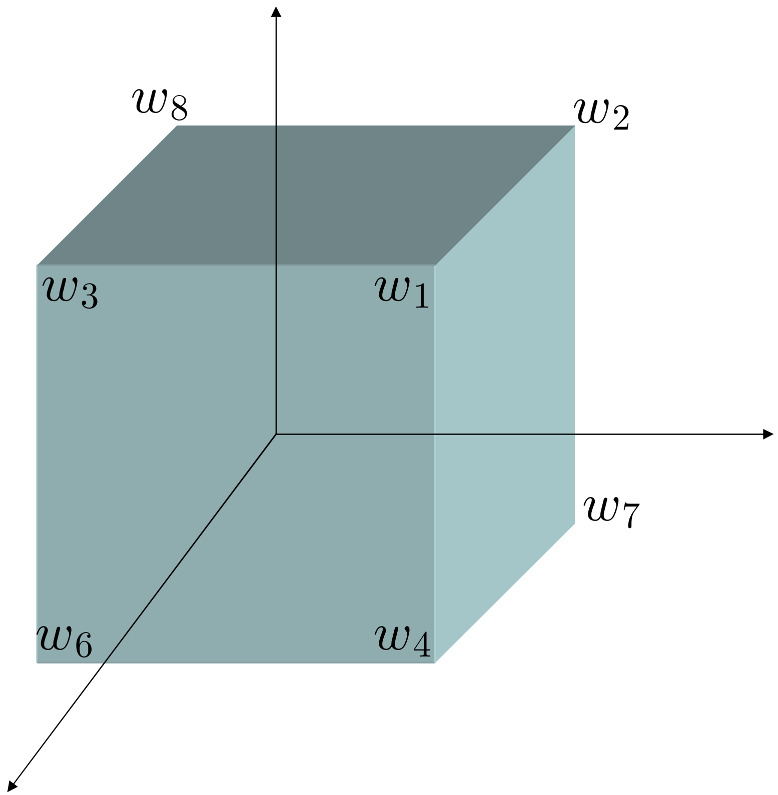
A.S. - Unit Cube

For linear controllers, need only consider the eight corners of the unit cube.



A.S. - Outputs

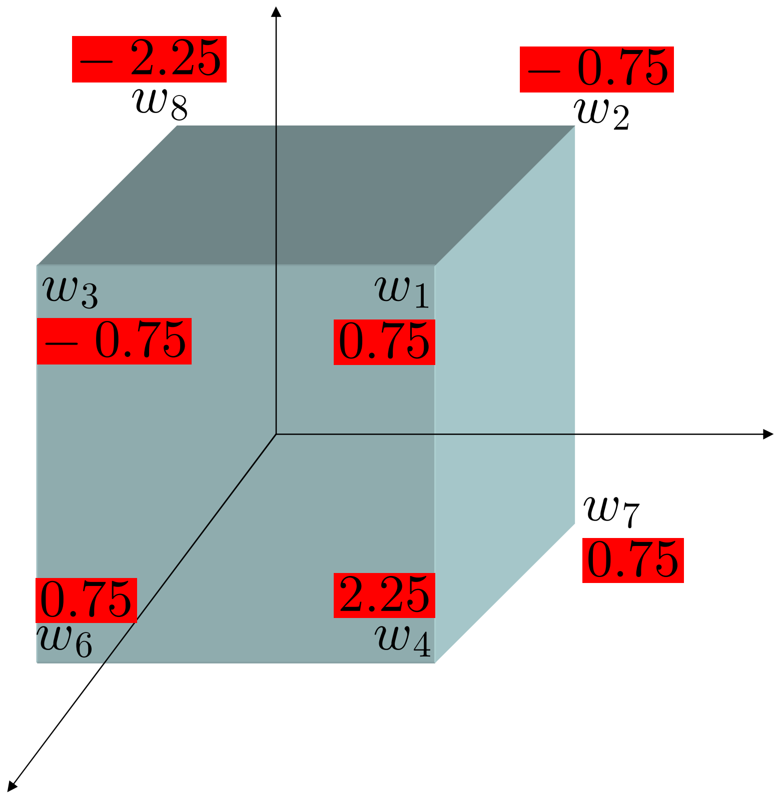
For linear controllers, need only consider the eight corners of the unit cube.



$$
\begin{aligned}
\|z_i\|_\infty &= 3 + |u_i| \quad \text{for } i = 1,2,3,5,6,7 \\
\|z_4\|_\infty &= \max\{|u_4 - 3|, |u_4 - 6|\} \\
\|z_8\|_\infty &= \max\{|u_8 + 3|, |u_8 + 6|\}
\end{aligned}
$$

A.S. - Best Linear

For linear controllers, need only consider the eight corners of the unit cube.

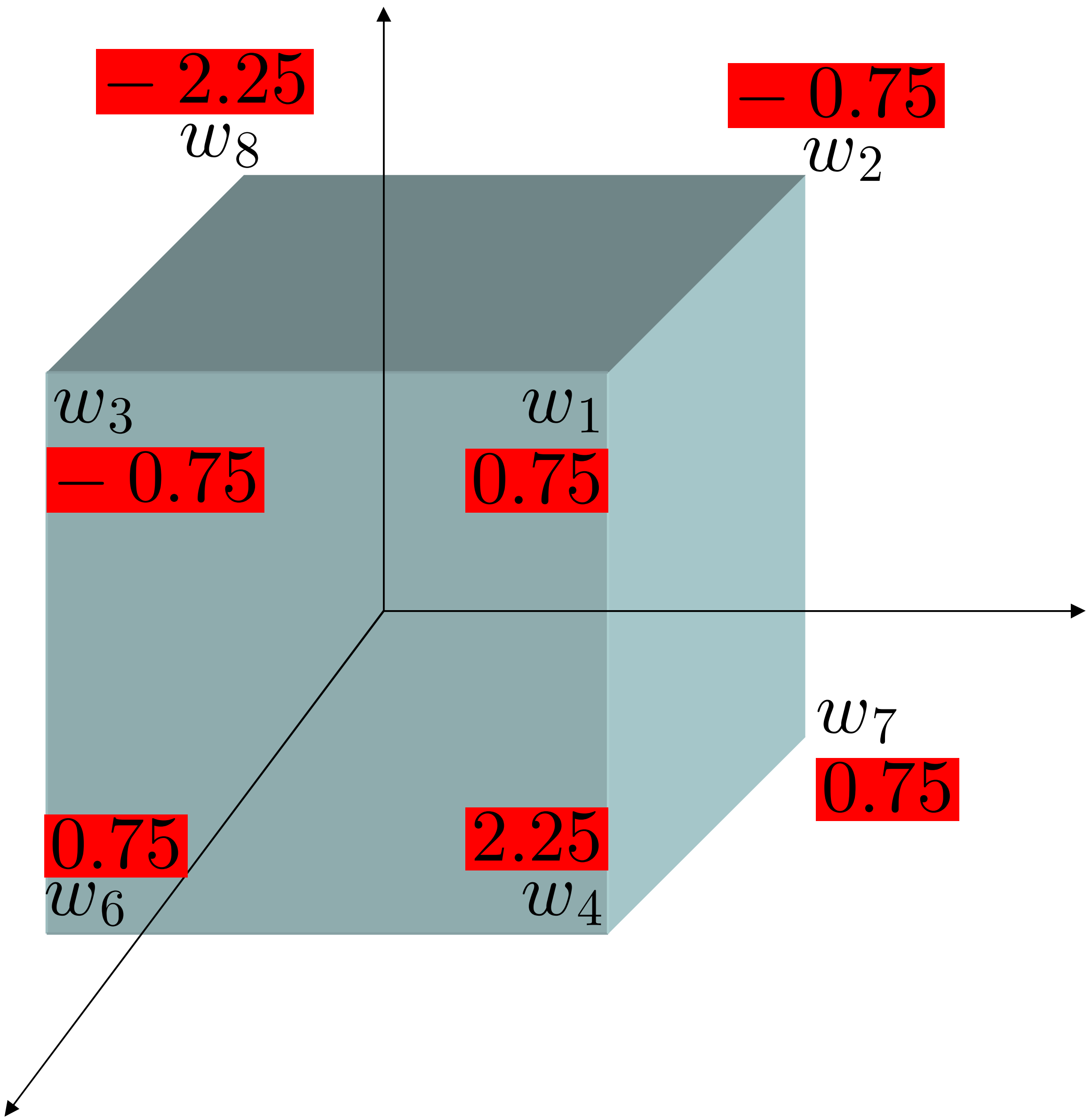


$$
\begin{aligned}
\|z_i\|_\infty &= 3 + |u_i| \quad \text{for } i = 1,2,3,5,6,7 \\
\|z_4\|_\infty &= \max\{|u_4 - 3|, |u_4 - 6|\} \\
\|z_8\|_\infty &= \max\{|u_8 + 3|, |u_8 + 6|\}
\end{aligned}
$$

Best linear controller achieves closed-loop norm of 3.75; achieved by

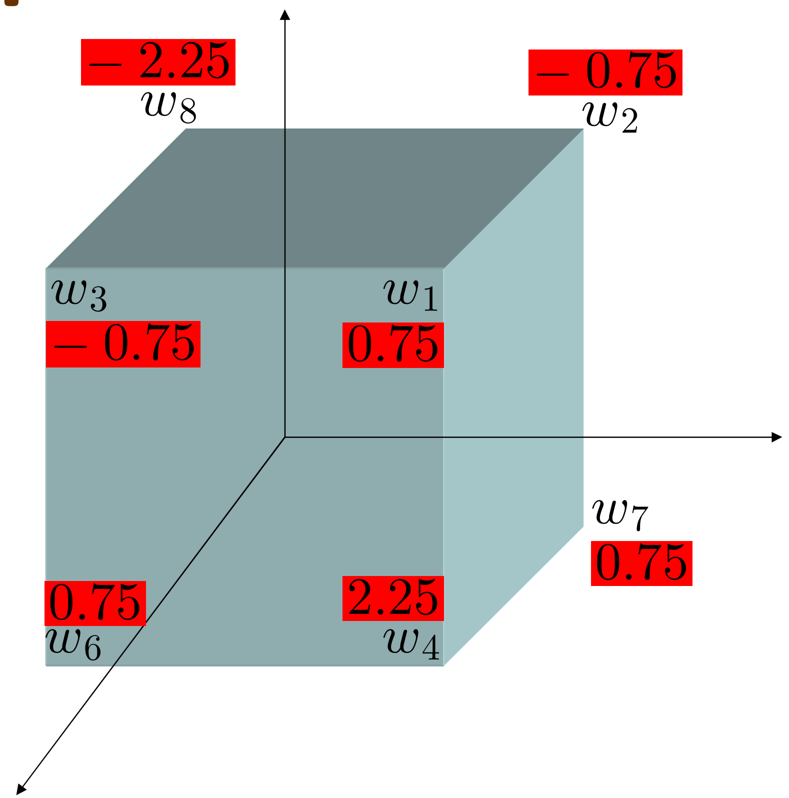
$$
u = \begin{bmatrix} 0.75 & -0.75 & 0.75 \end{bmatrix} w
$$

A.S. - Linear Controller



A.S. - Nonlinear Control

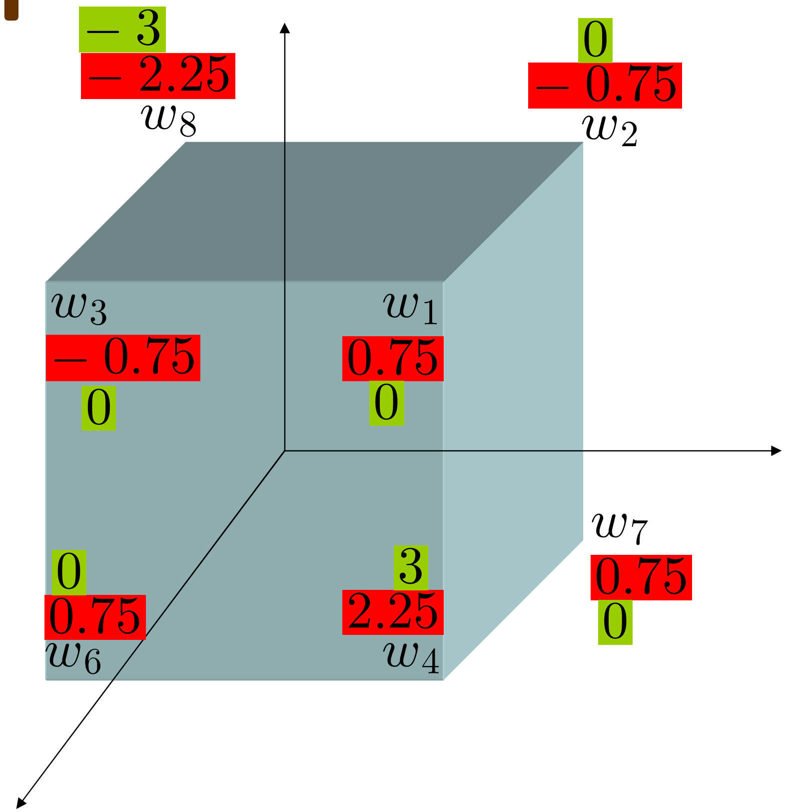
For nonlinear controllers, also consider the eight corners of the unit cube.



$$
\begin{aligned}
\|z_i\|_\infty &= 3 + |u_i| \quad \text{for } i = 1,2,3,5,6,7 \\
\|z_4\|_\infty &= \max\{|u_4 - 3|, |u_4 - 6|\} \\
\|z_8\|_\infty &= \max\{|u_8 + 3|, |u_8 + 6|\}
\end{aligned}
$$

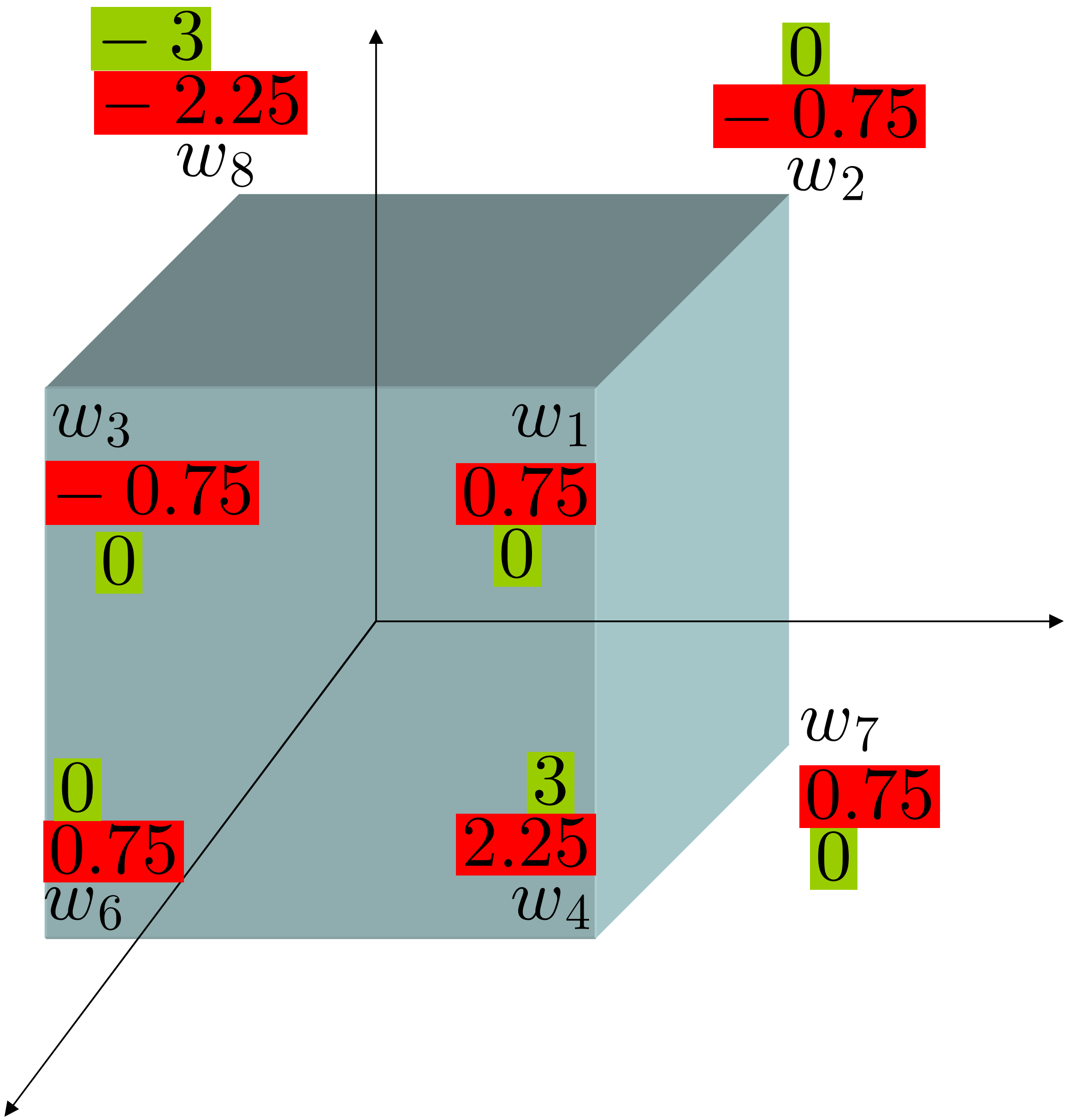
A.S. - Nonlinear Control

For nonlinear controllers, also consider the eight corners of the unit cube.



$$
\begin{aligned}
\|z_i\|_\infty &= 3 + |u_i| \quad \text{for } i = 1,2,3,5,6,7 \\
\|z_4\|_\infty &= \max\{|u_4 - 3|, |u_4 - 6|\} \\
\|z_8\|_\infty &= \max\{|u_8 + 3|, |u_8 + 6|\}
\end{aligned}
$$

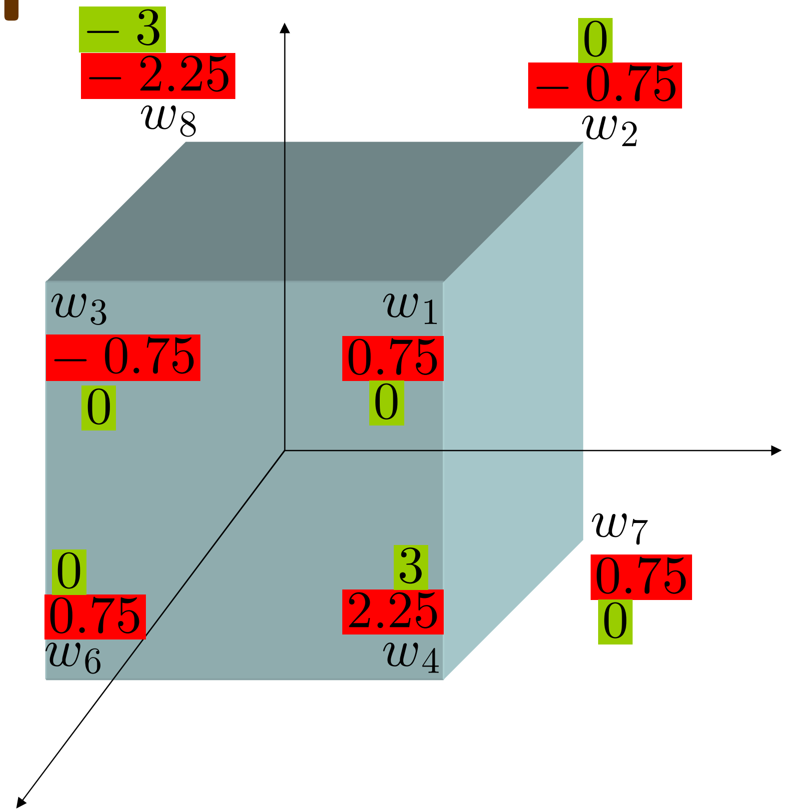
A.S. - Nonlinear Control



A.S. - Nonlinear Control

For nonlinear controllers, also consider the eight corners of the unit cube.

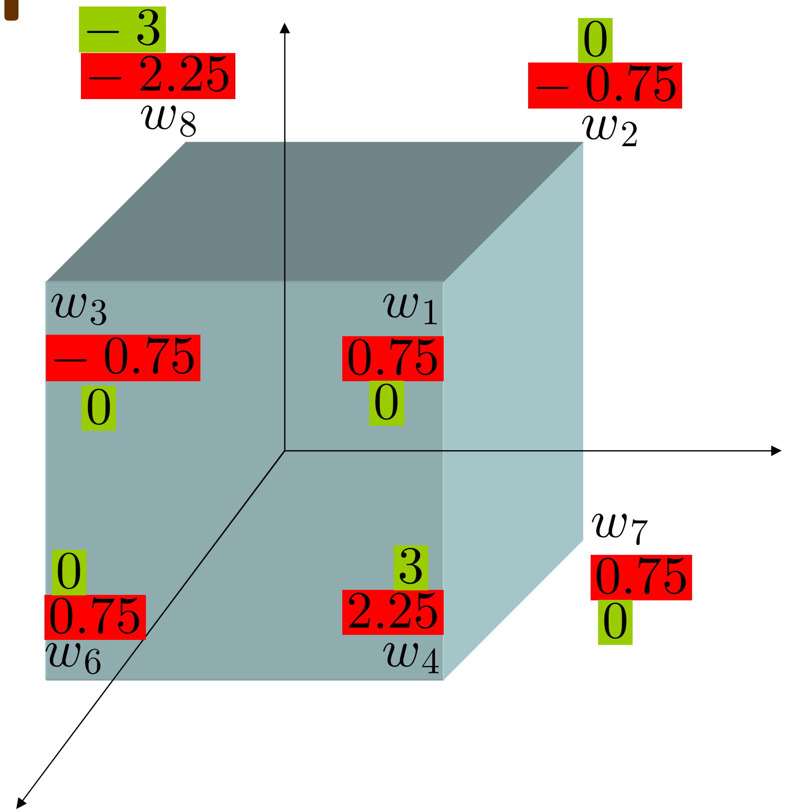
$$\|z_i\|_\infty = 3 + |u_i| \quad \text{for } i = 1, 2, 3, 5, 6, 7$$
$$\|z_4\|_\infty = \max\{|u_4 - 3|, |u_4 - 6|\}$$
$$\|z_8\|_\infty = \max\{|u_8 + 3|, |u_8 + 6|\}$$



Best nonlinear controller achieves closed-loop norm of 3.00.

A.S. - Nonlinear Control

For nonlinear controllers, also consider the eight corners of the unit cube.

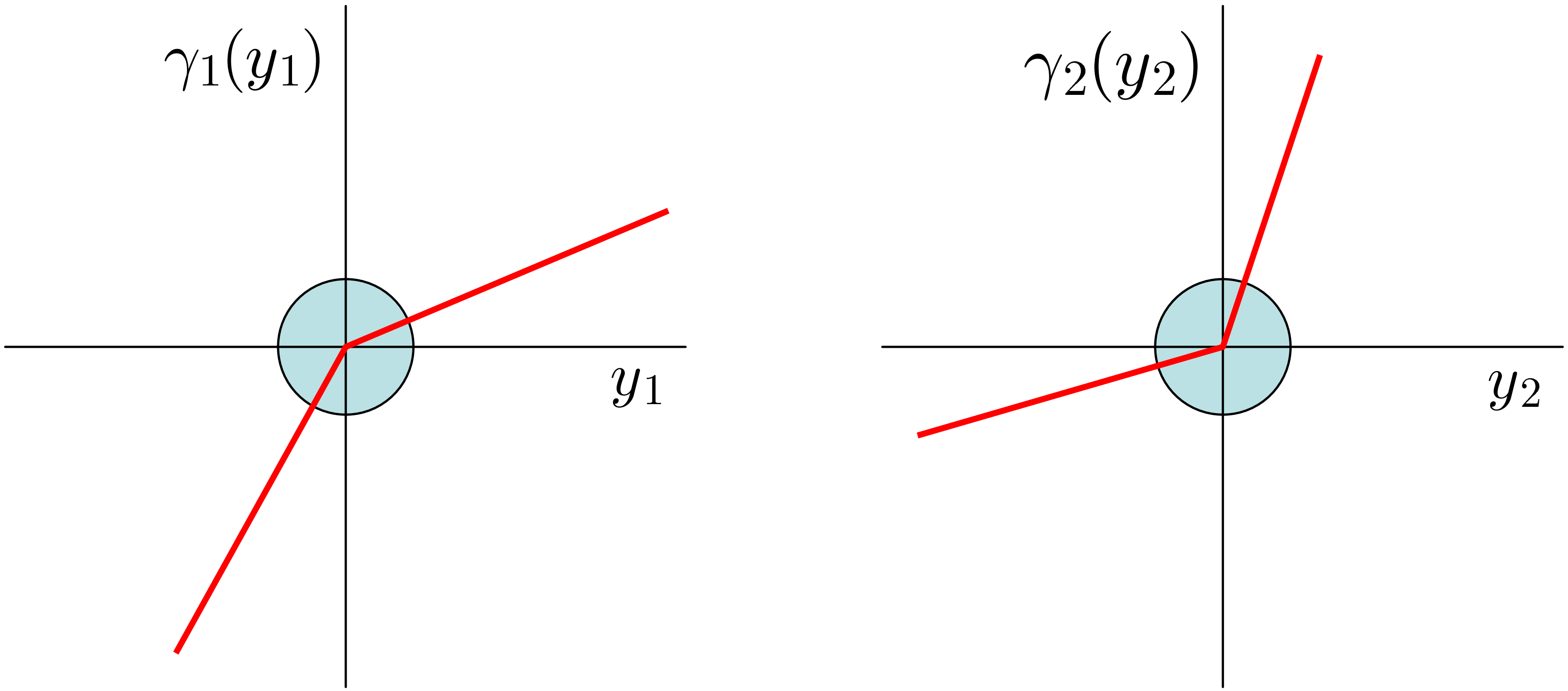


$$\begin{aligned} \|z_i\|_\infty &= 3 + |u_i| \quad \text{for } i = 1,2,3,5,6,7 \\ \|z_4\|_\infty &= \max\{|u_4 - 3|, |u_4 - 6|\} \\ \|z_8\|_\infty &= \max\{|u_8 + 3|, |u_8 + 6|\} \end{aligned}$$

Best nonlinear controller achieves closed-loop norm of 3.00.

- Nonlinear outperforms linear - for a centralized problem!

Back to Witsenhausen



Consider controllers of the form

$$u_1 = \begin{cases} a_- y_1 & \text{if } y_1 \leq 0 \\ a_+ y_1 & \text{if } y_1 \geq 0 \end{cases} \qquad u_2 = \begin{cases} b_- y_2 & \text{if } y_2 \leq 0 \\ b_+ y_2 & \text{if } y_2 \geq 0 \end{cases}$$

Notation

- $Q(\cdot)$

  ratio $\frac{\|z\|_2}{\|w\|_2}$ for a given controller and given noise w_1, w_2

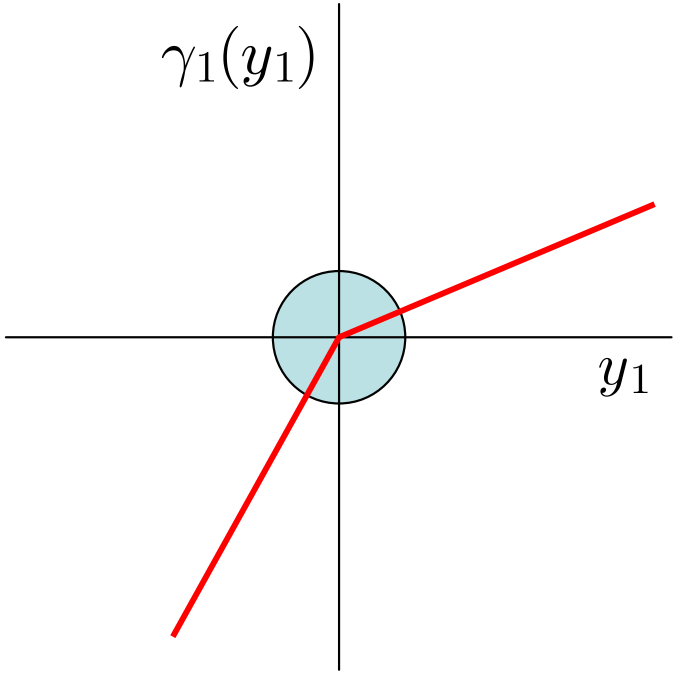
- $J(\cdot)$

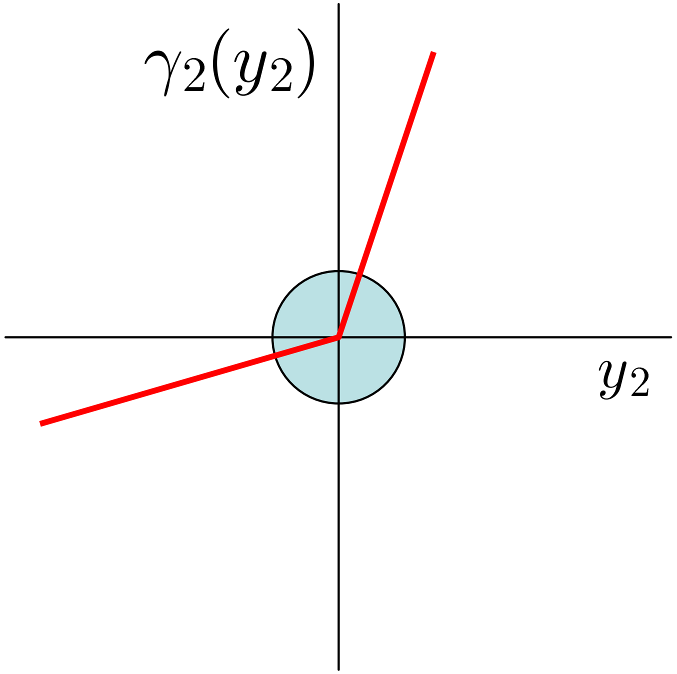
  worst case ratio $\sup_{w \neq 0} \frac{\|z\|_2}{\|w\|_2}$ for a given controller, i.e. the cost associated with that controller

- J^*

  best achievable cost for a given class of controllers

Back to Witsenhausen

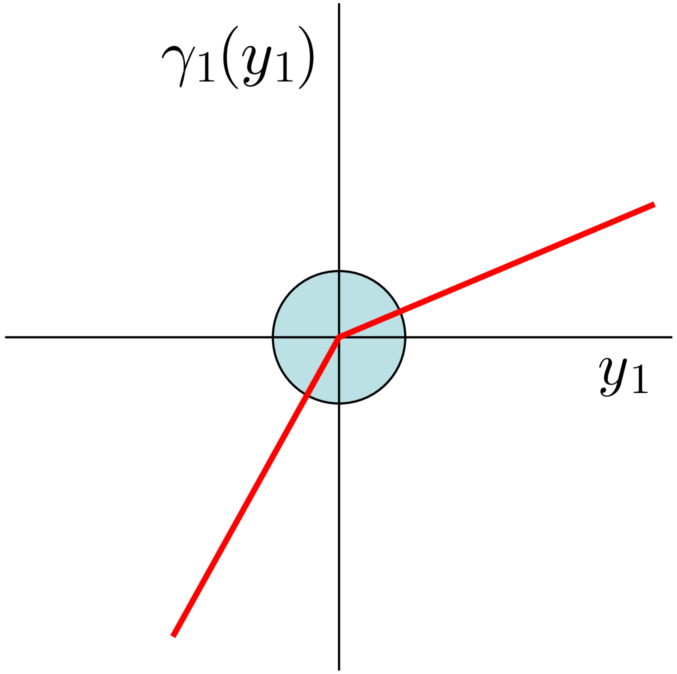


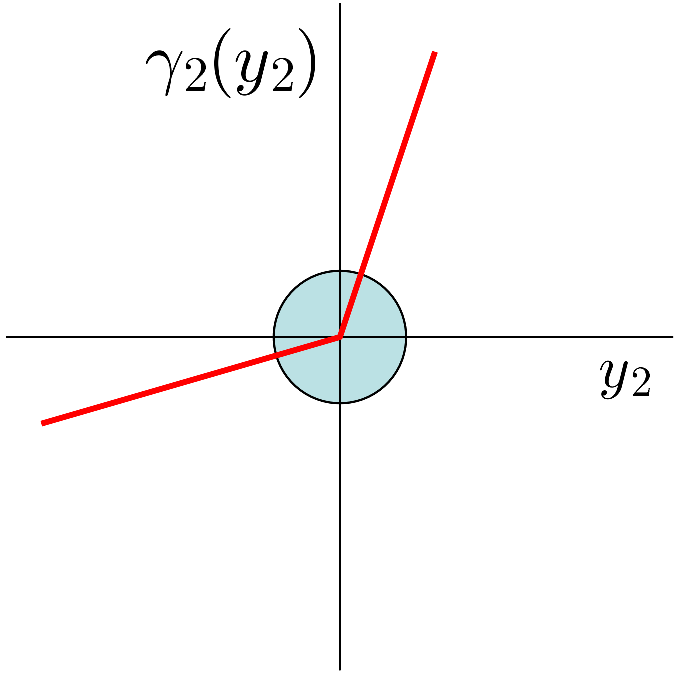


Find best piecewise affine controller

$$J_{pa}^* = \inf_{a_-,a_+} \inf_{b_-,b_+} J_{pa}(a_-, a_+, b_-, b_+)$$

Piecewise Affine Cost





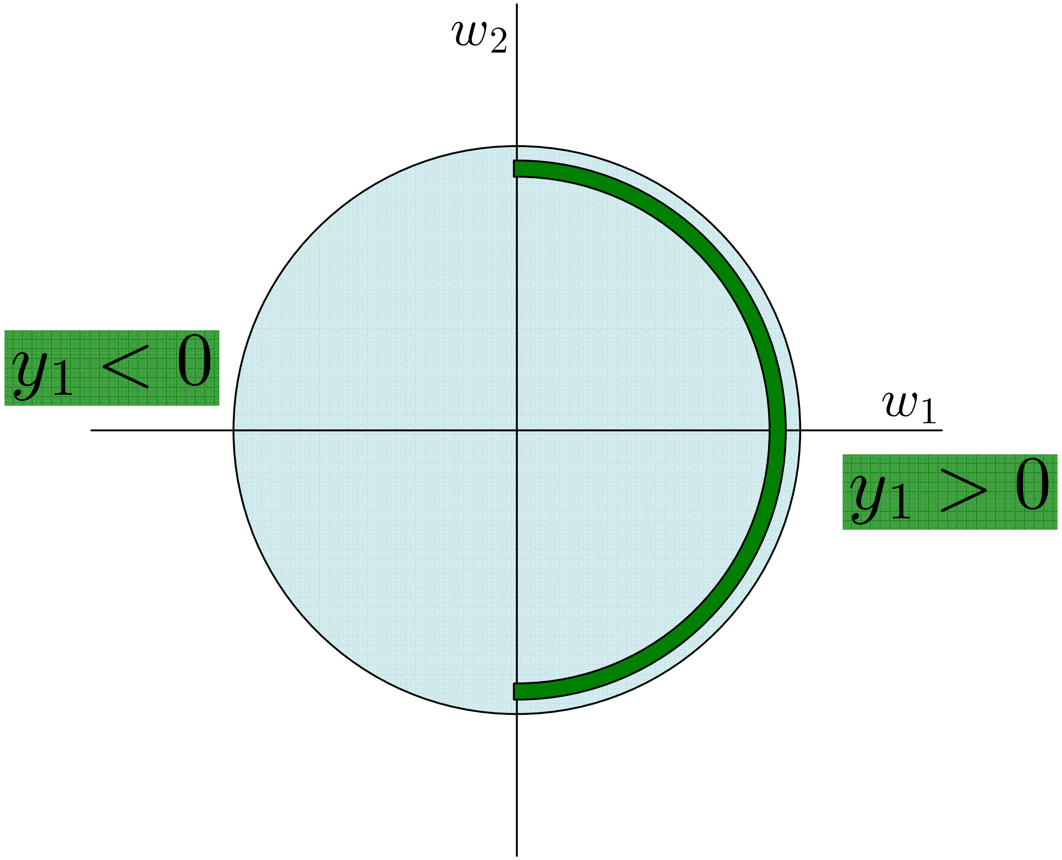
Find best piecewise affine controller

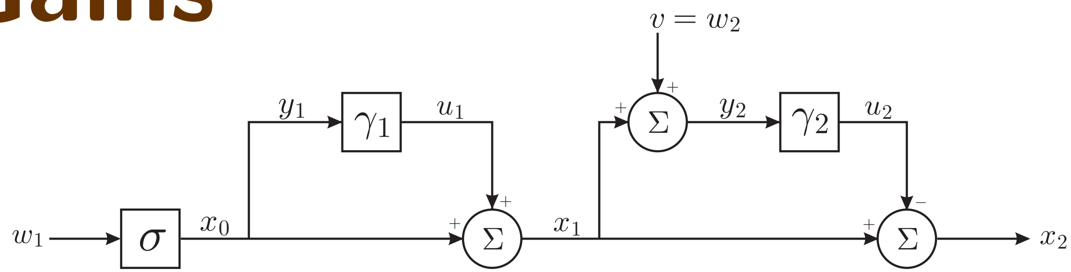
$$J^*_{pa} \; = \; \inf_{a_-,a_+} \inf_{b_-,b_+} J_{pa}(a_-,a_+,b_-,b_+)$$

As with linear, only the direction of the noise matters

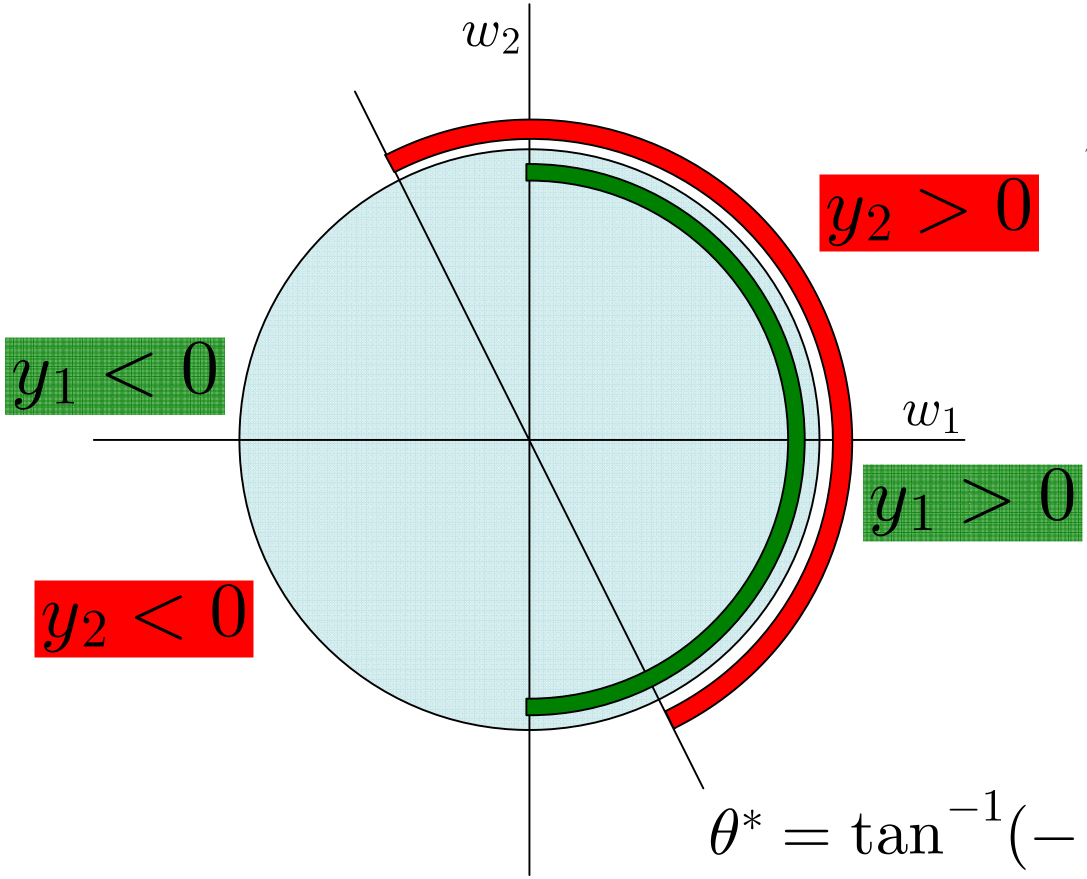
$$J_{pa}(a_-,a_+,b_-,b_+) \; = \; \sup_{\theta} \; Q_{pa}(a_-,a_+,b_-,b_+,\theta)$$

Regions of Different Gains

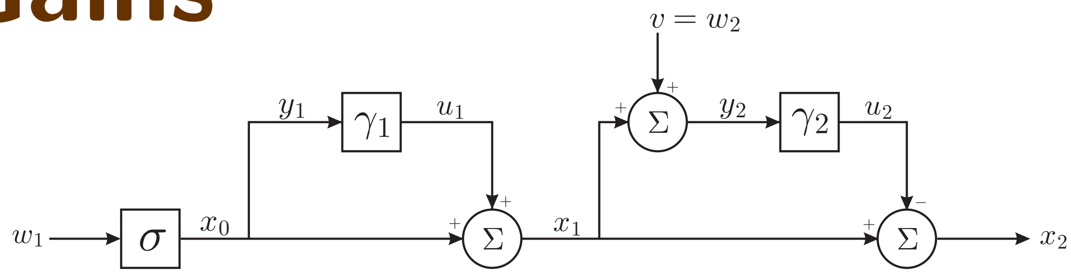




Regions of Different Gains

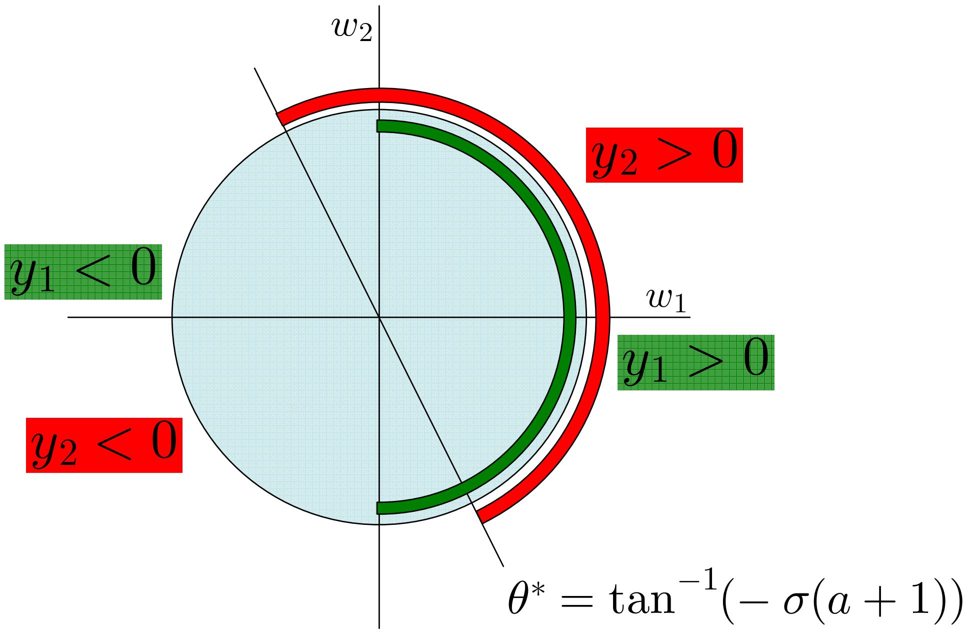


$$\theta^* = \tan^{-1}(-\sigma(a+1))$$



$$J_{pa}(a_-, a_+, b_-, b_+) \;=\; \sup_{\theta} \; Q_{pa}(a_-, a_+, b_-, b_+, \theta)$$

Regions of Different Gains



At θ^*, we have $y_2 = 0$. Thus

$$\sup_{\theta} Q_{pa}(a, b_-, b_+, \theta) \;\geq\; Q_{pa}(a, b_-, b_+, \theta^*) \quad \text{for all } b_-, b_+$$

and so

$$J^*_{\mathsf{pa}}(a) \;=\; \inf_{b_-,b_+} \sup_{\theta} Q_{pa}(a, b_-, b_+, \theta) \;\geq\; Q_{pa}(a, b_-, b_+, \theta^*)$$

Set Second Controller

- We'd like

$$\left.\frac{\partial}{\partial\theta} Q_{pa}(a, b_-, b_+, \theta)\right|_{\theta=\theta^*} = 0$$

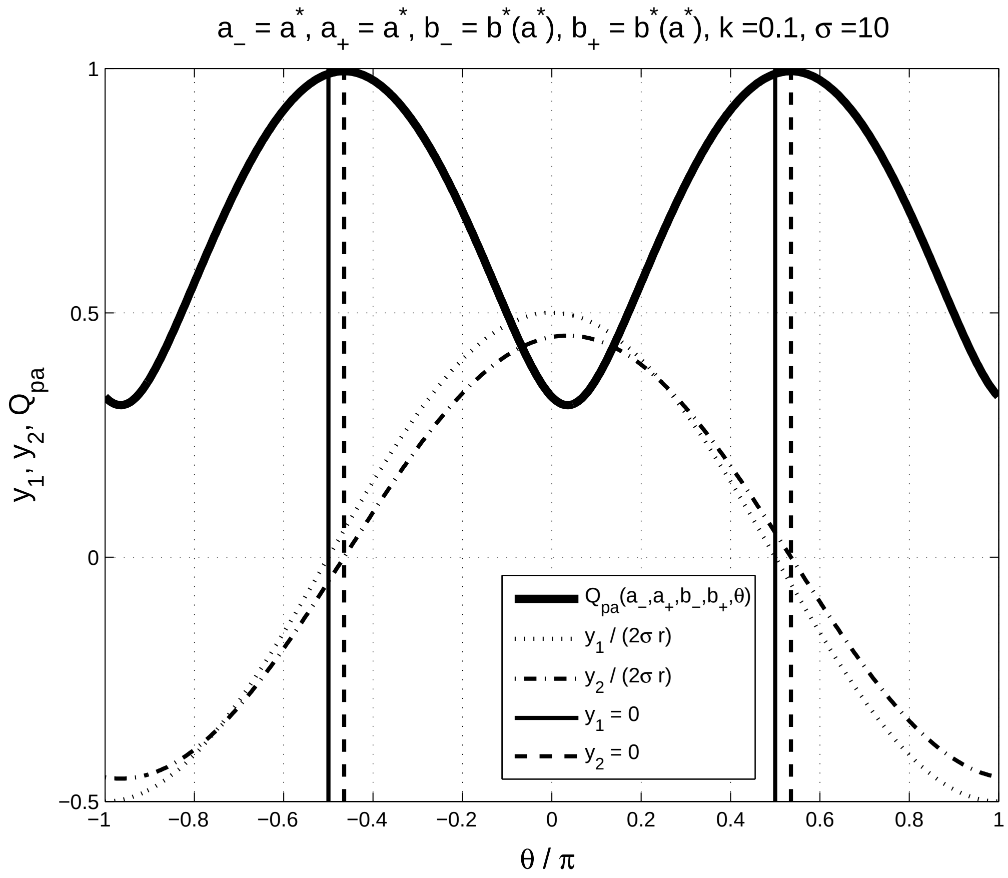
- and also

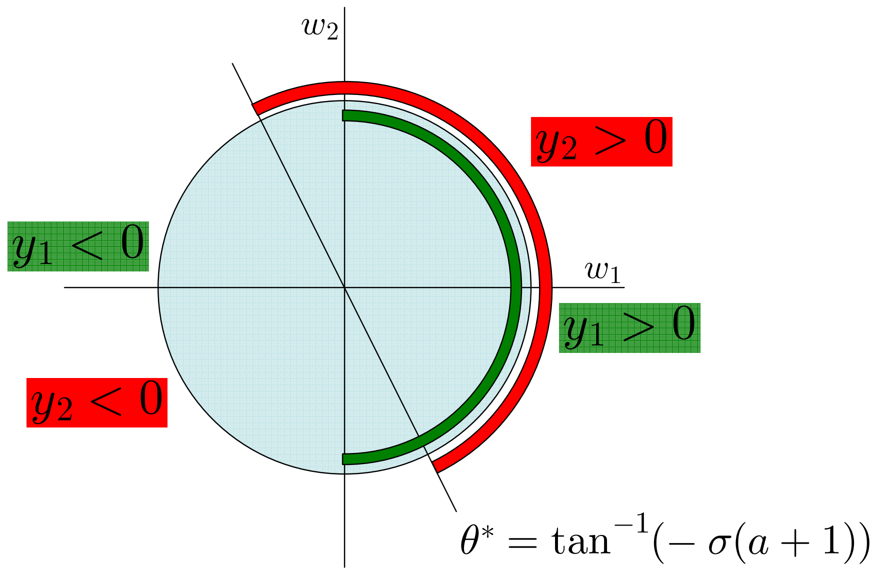
$$\left.\frac{\partial^2 Q_{pa}^2(a, b_-, b_+, \theta)}{\partial\theta^2}\right|_{\theta=\theta^*} < 0$$

- Achieved when we choose $b_- = b_+ = b^*(a)$ where

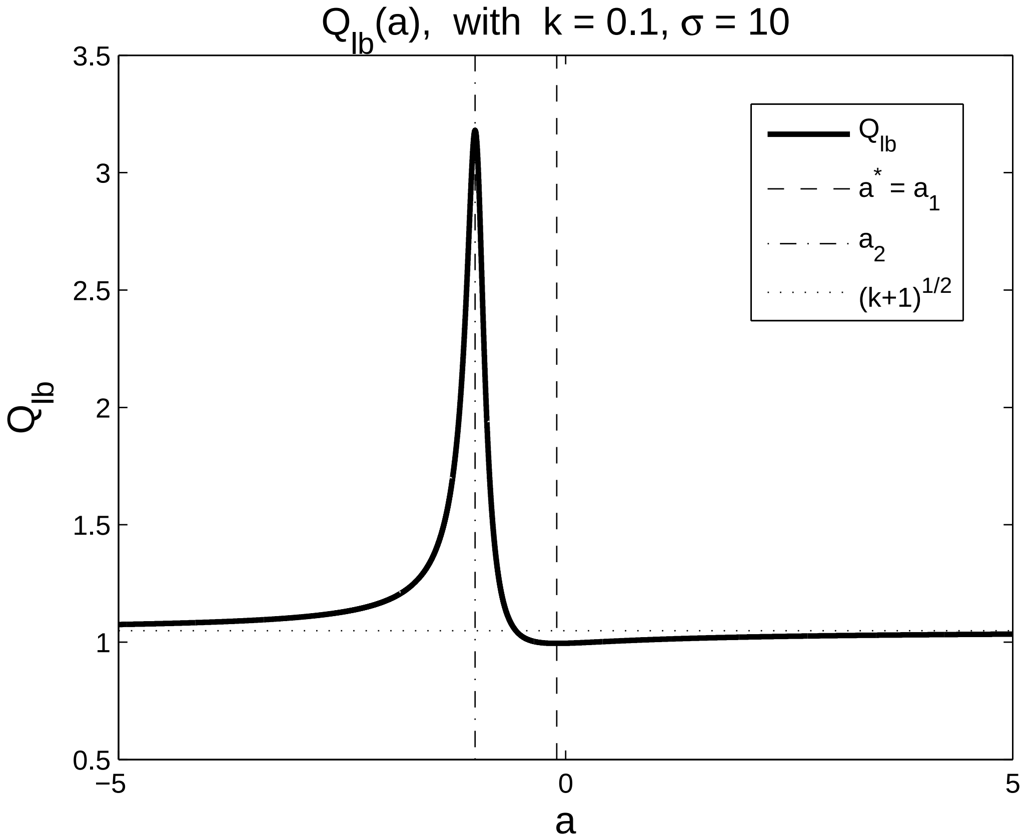
$$b^*(a) = \frac{\sigma^2(k^2a^2 + (1+a)^2)}{1 + \sigma^2(1+a)^2}$$

Gain vs. Noise for Optimal Controllers





Cost vs. First Controller



$$a^* = 2\left(-\alpha + \sqrt{\alpha^2 - \beta^2}\right)/\beta^2$$

where $\alpha = k^2\sigma^2 + k^2 + 1$ and $\beta = 2\sigma k$.

General Nonlinear Controller

Given any $\gamma_1, \gamma_2 : \mathbb{R} \to \mathbb{R}$, let

$$a = \gamma_1(1)$$

and then let

$$\theta_0 \;=\; \theta^*(a)$$

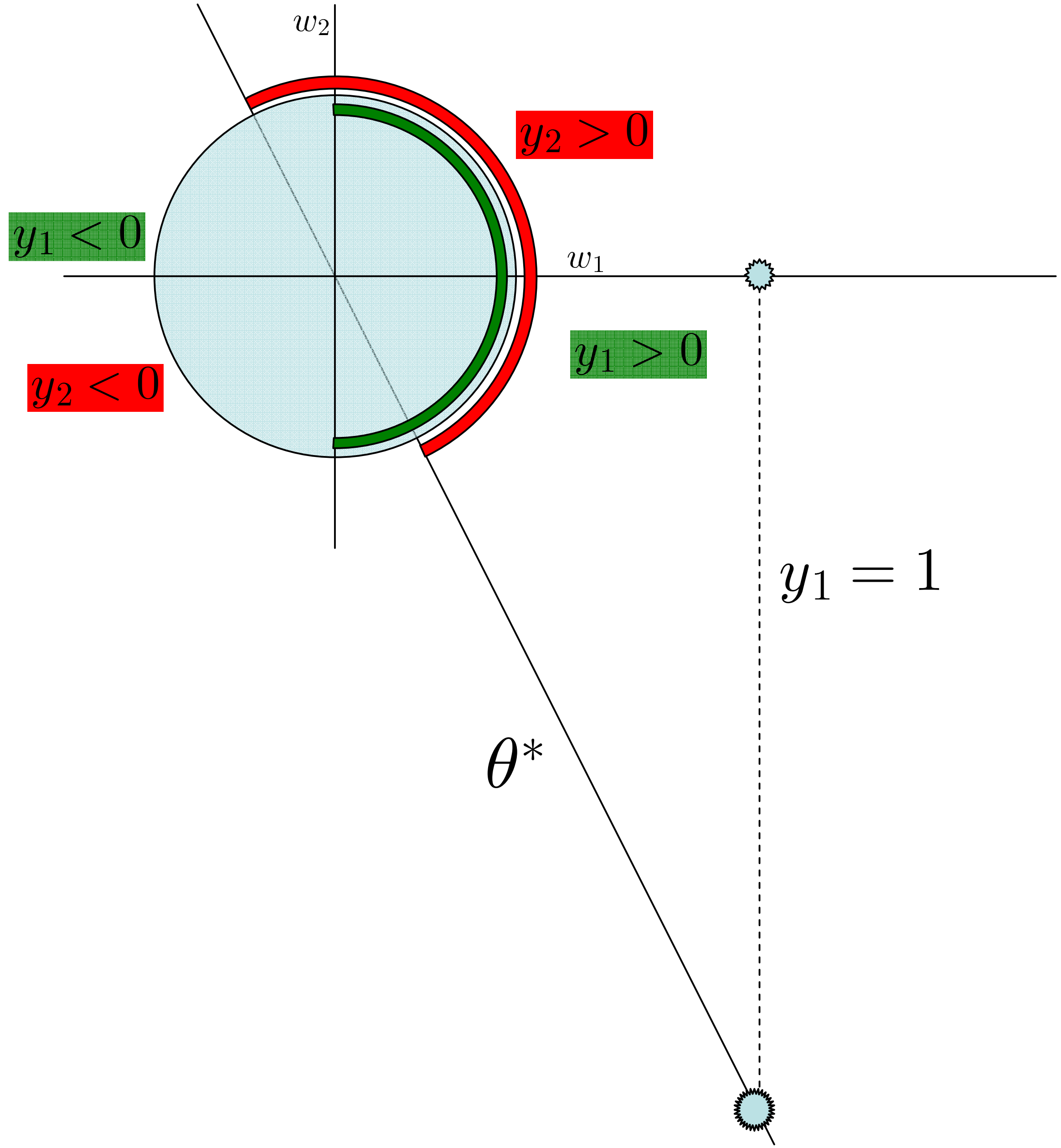
and further choose r_0 such that

$$y_1(r_0, \theta_0) \;=\; 1$$

and

$$y_2(r_0, \theta_0) \;=\; 0$$

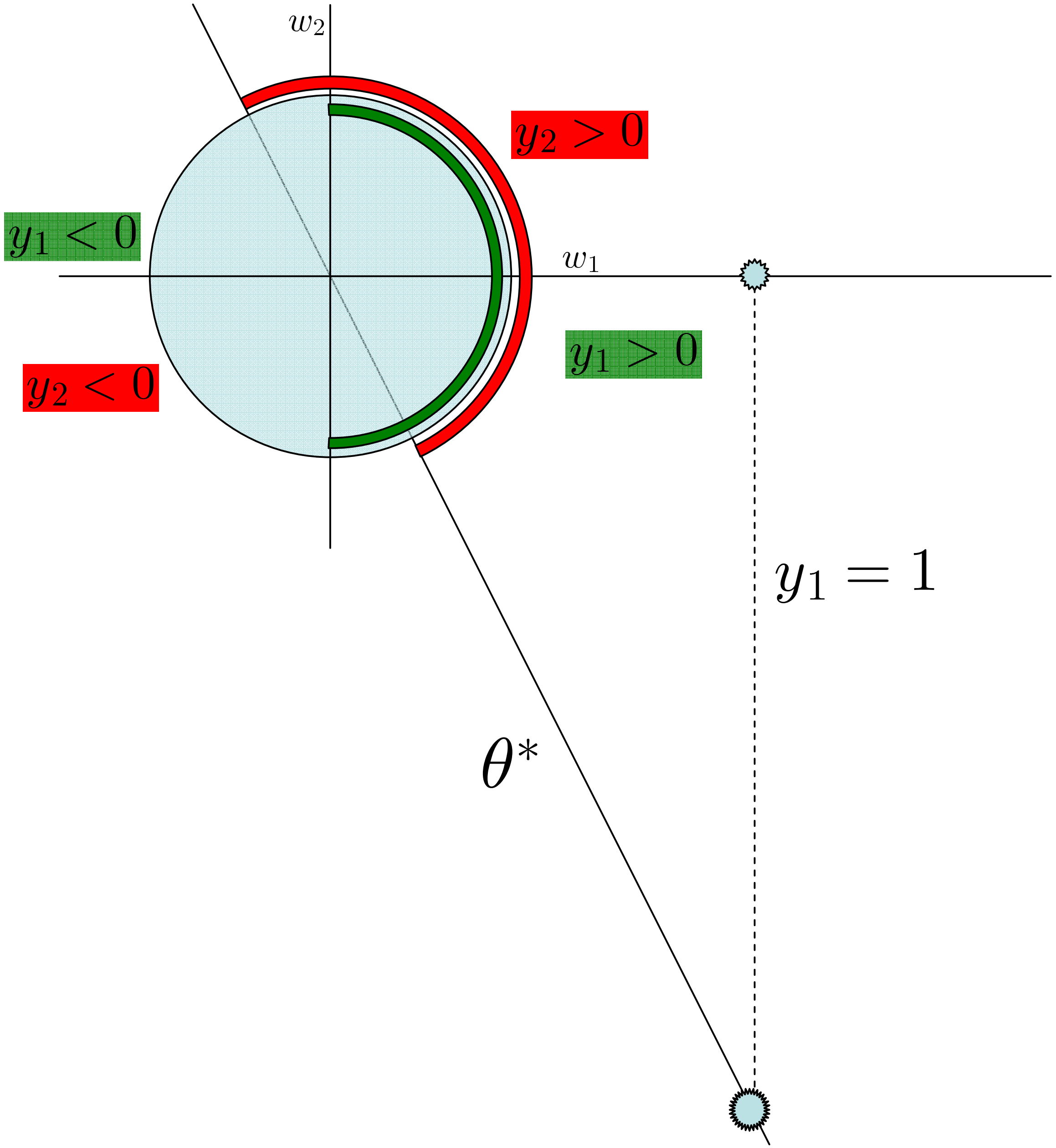
General Nonlinear Controller



General Nonlinear Controller

Obvious that

$$J^* \;\leq\; J_l^*$$



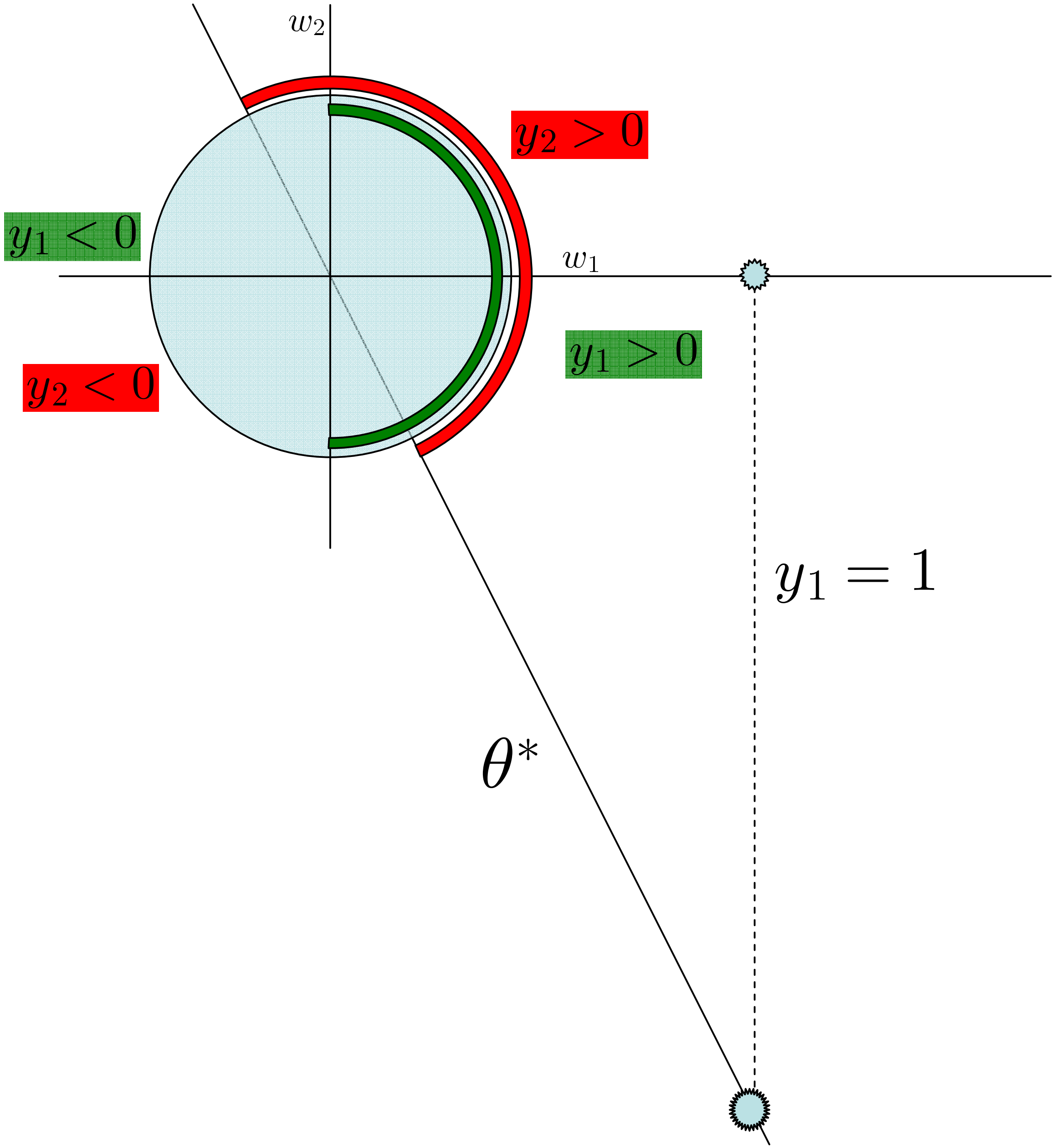
General Nonlinear Controller

Obvious that

$$J^* \leq J_l^*$$

Now

$$Q(\gamma_1, \gamma_2, r_0, \theta_0)$$



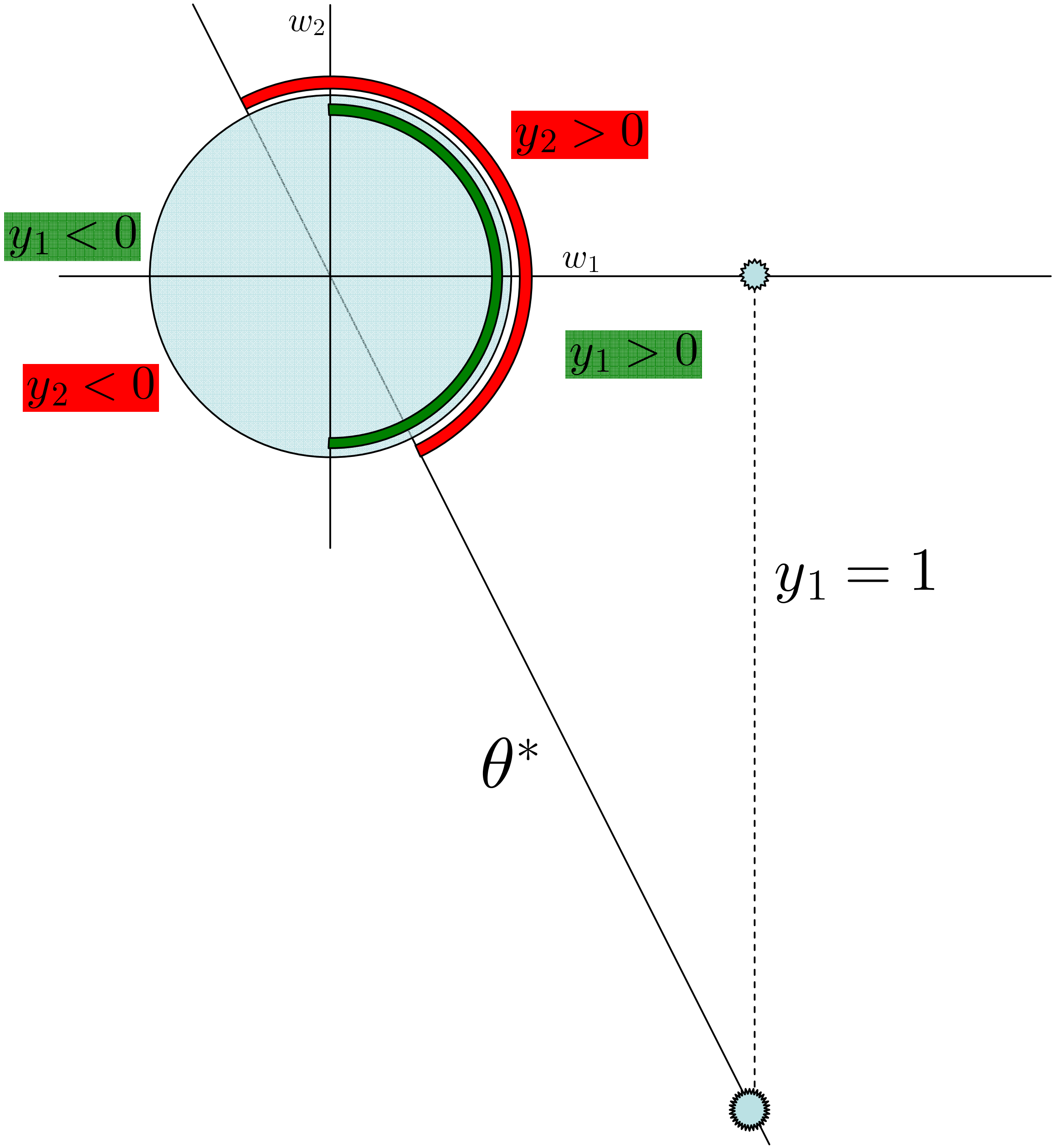
General Nonlinear Controller

Obvious that

$$J^* \;\leq\; J_l^*$$

Now

$$\begin{aligned} &Q(\gamma_1, \gamma_2, r_0, \theta_0) \\ &\quad = \; Q_l(a, b^*(a), \theta^*(a)) \end{aligned}$$



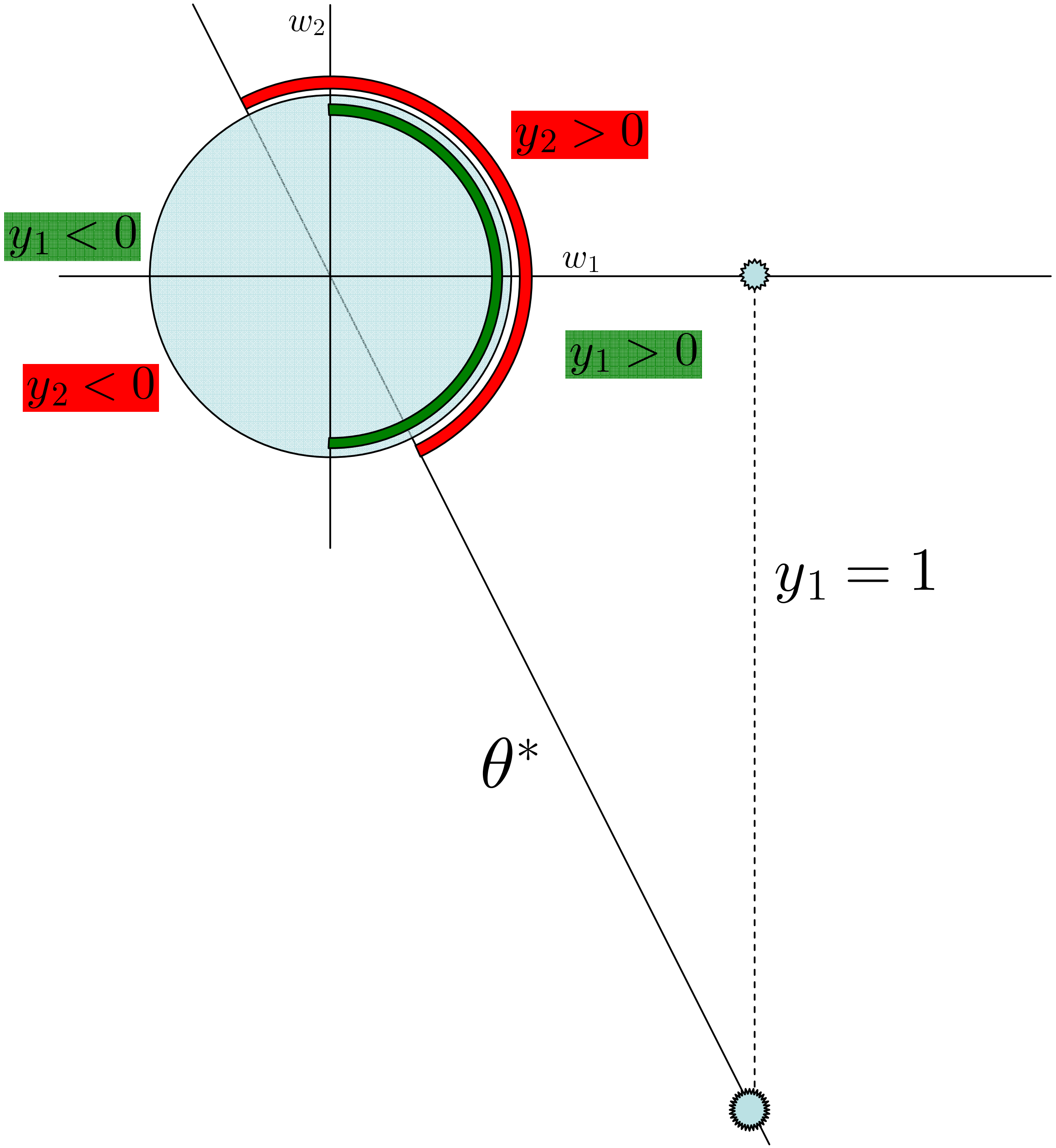
General Nonlinear Controller

Obvious that

$$J^* \leq J_l^*$$

Now

$$\begin{aligned} &Q(\gamma_1, \gamma_2, r_0, \theta_0) \\ &\quad = Q_l(a, b^*(a), \theta^*(a)) \\ &\quad = J_l(a, b^*(a)) \end{aligned}$$



General Nonlinear Controller

Obvious that

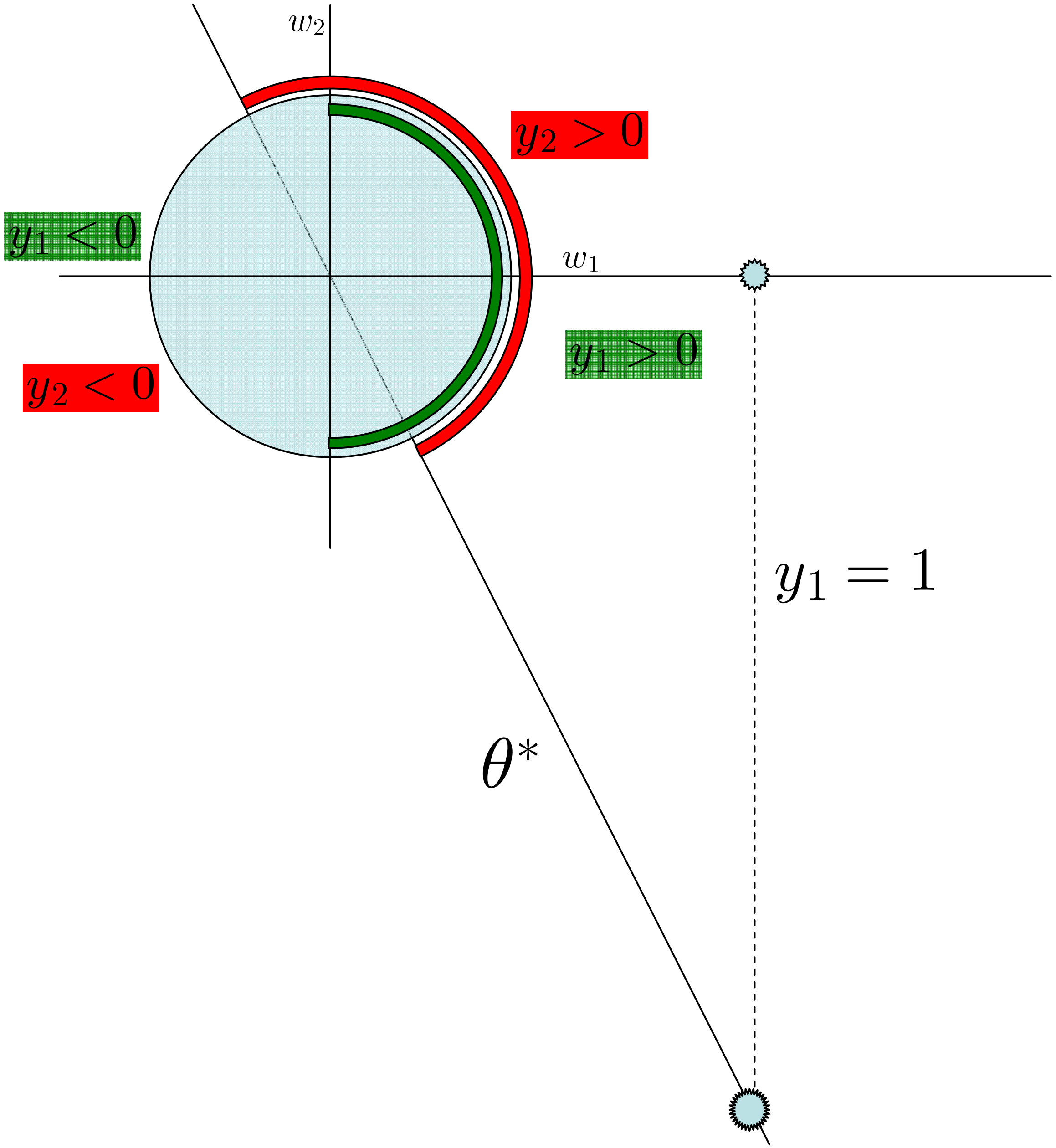
$$J^* \;\leq\; J_l^*$$

Now

$$\begin{aligned} Q(\gamma_1, \gamma_2, r_0, \theta_0) \\ &= Q_l(a, b^*(a), \theta^*(a)) \\ &= J_l(a, b^*(a)) \end{aligned}$$

which yields

$$J(\gamma_1, \gamma_2) \;\geq\; J_l(a, b^*(a))$$



General Nonlinear Controller

Obvious that

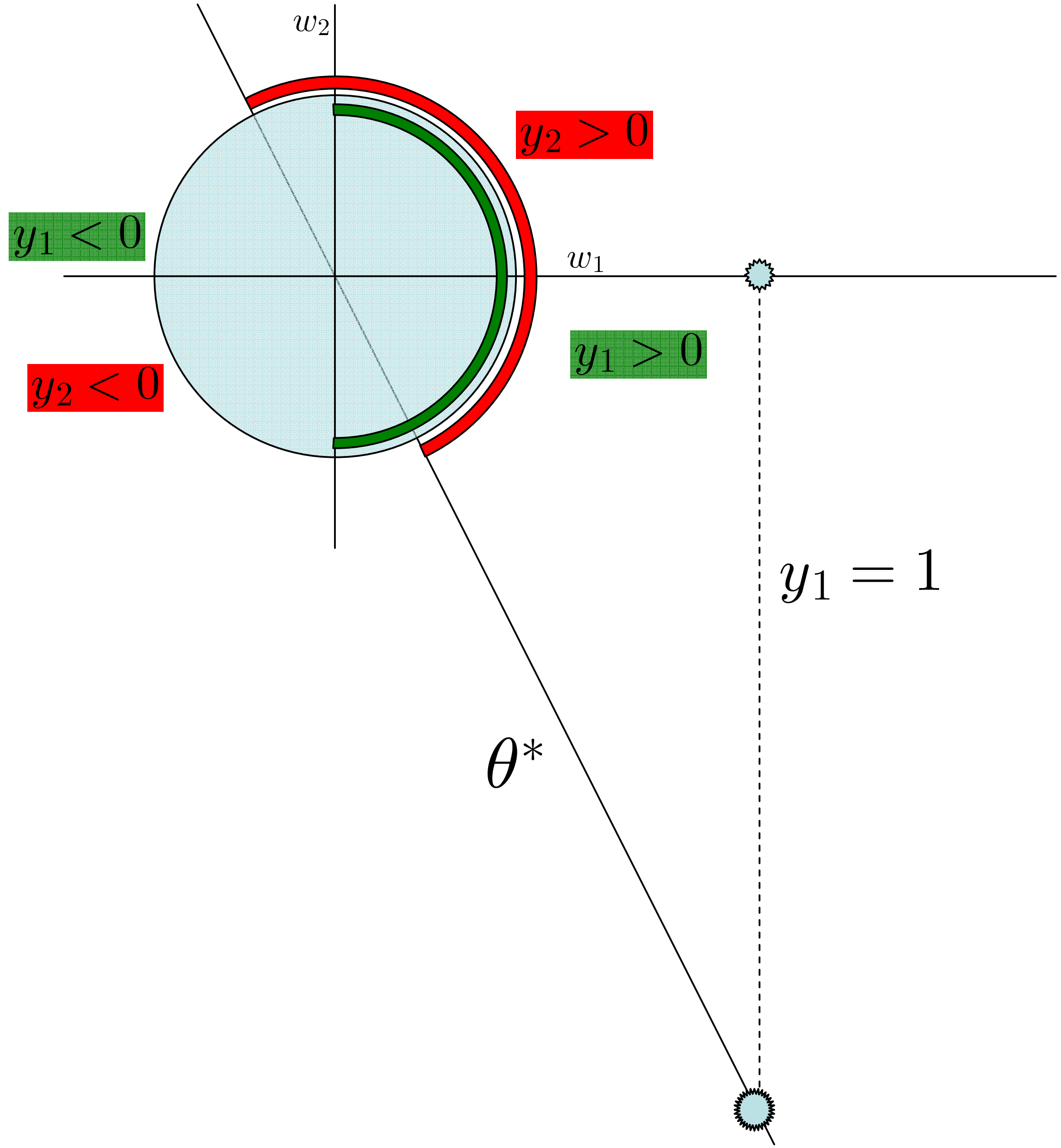
$$J^* \leq J_l^*$$

and we've now shown that

$$J^* \geq J_l^*$$

thus

$$J^* = J_l^*$$



Redundancy

- Argued that piecewise affine dominates general nonlinear.
- Showed that best piecewise affine controller is linear.
- Showed that linear dominates general nonlinear.

Redundancy

- Argued that piecewise affine dominates general nonlinear.
- Showed that best piecewise affine controller is linear.
- Showed that linear dominates general nonlinear.

Why?

- Generalization

Redundancy

- Argued that piecewise affine dominates general nonlinear.
- Showed that best piecewise affine controller is linear.
- Showed that linear dominates general nonlinear.

Why?

- Generalization
- Nobody believed me!

Related Work

- Separation / Linear Optimality over Lossy Networks
  - X. Liu, A. Goldsmith
  - L. Schenato, B. Sinopoli, M. Franceschetti, K. Poola, S. Sastry
- (Decentralized) Robust Control
  - Ian Petersen, Valery Ugrinovskii

Conclusion(s)

- Very simple question gives rise to extremely rich structure.
- Linear controllers are uniformly optimal for the Witsenhausen Counterexample.
- Many new surprises await!

System Norm	Induced by	Linear Optimal for Centralized?	Linear Optimal for Decentralized?
$\|\cdot\|_1$	$\frac{\|\cdot\|_\infty}{\|\cdot\|_\infty}$	No	??
$\|\cdot\|_2$ / LQG	-	Yes	Sometimes (PN/QI)
$\|\cdot\|_\infty$	$\frac{\|\cdot\|_2}{\|\cdot\|_2}$	Yes	??