Non-Neutral Decision Making in Stochastic Teams and Games

TAMER BAŞAR
Beckman Institute
Dept ECE, CAS, CSL and ITI, UIUC
basar1@illinois.edu

GAM Workshop
University of Warwick, Coventry
April 14-17, 2010

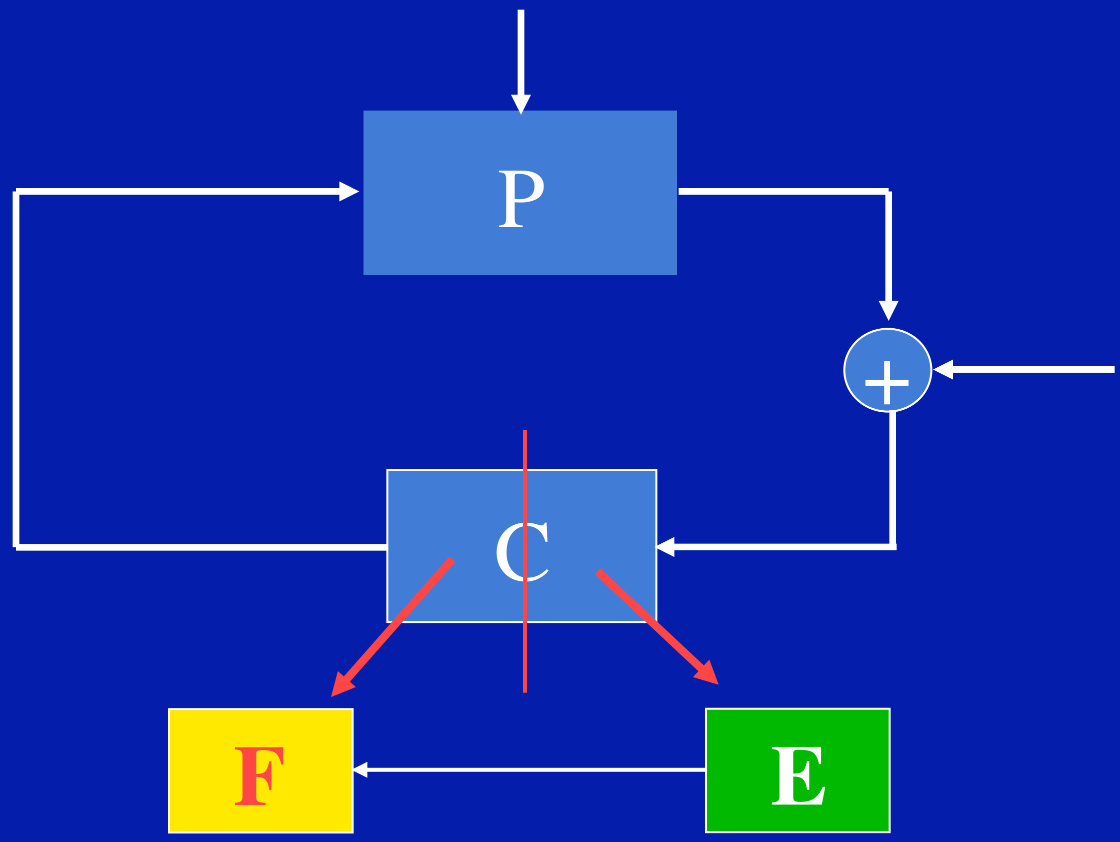
OUTLINE

- **Neutrality and non-classical information in control and dynamic games**
- **Some caveats and counter-examples**
- **Tractable problems with non-classical information**
- **Limited action teams / games**
- **Subtleties in games with noisy information channels (even with classical information)**
- **Conclusions**

Neutrality

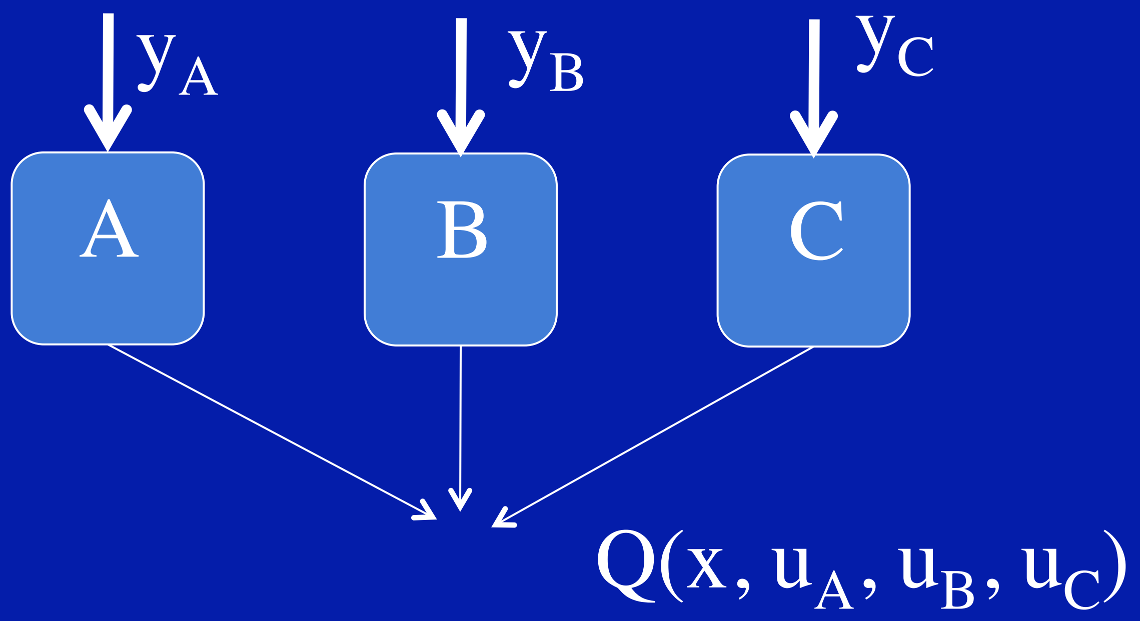
A stochastic control problem is neutral if, roughly speaking, the *quality* of information carried to future stages is independent of past controls. If control policies can shape future information, then problem is non-neutral. In this case, there is generally a conflict between *action* and *probing* roles of control -- dual control.

Separation / Neutrality

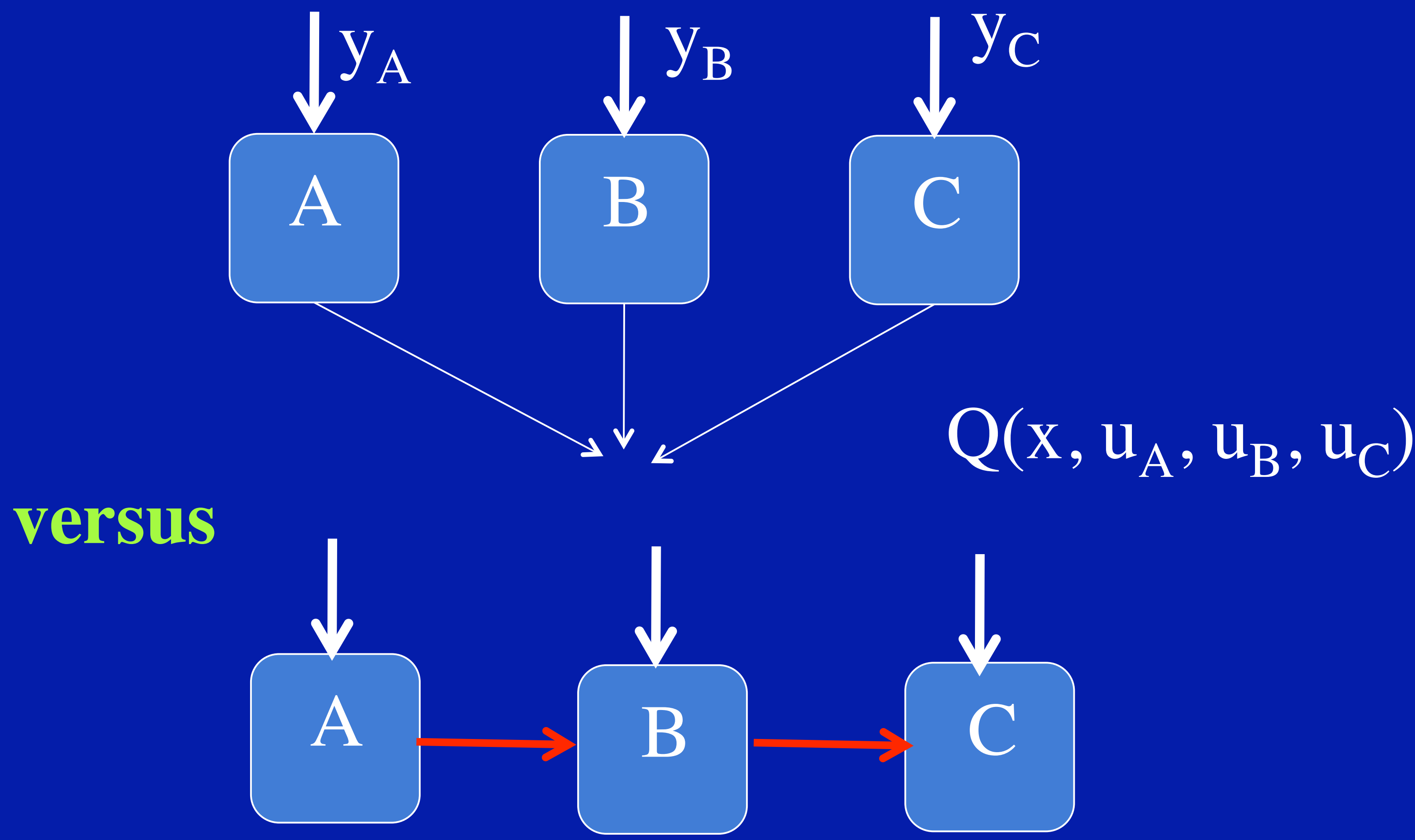


A stochastic decision problem is one with *non-classical information*, if a decision unit, **B**, that *follows* another one, **A**, and *whose actions are coupled*, does not have all the information acquired and used by **A**.

A stochastic decision problem is one with *non-classical information*, if a decision unit, **B**, that *follows* another one, **A**, and *whose actions are coupled*, does not have all the information acquired and used by **A**.

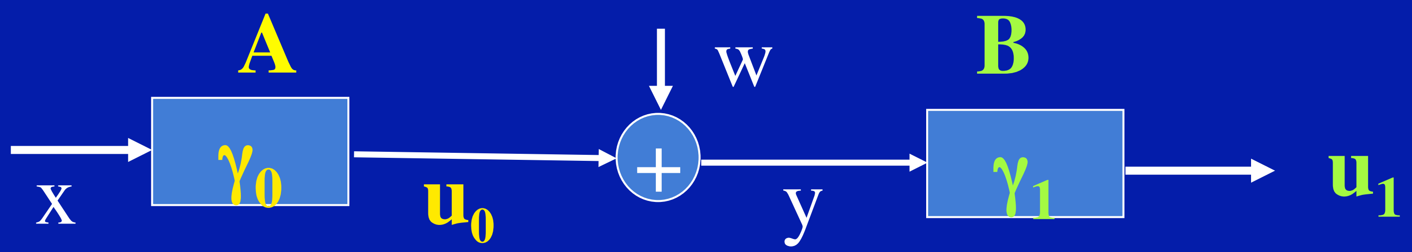


y_A
y_B
y_C
A
B
C
$Q(x, u_A, u_B, u_C)$



y_A
y_B
y_C
A
B
C
$Q(x, u_A, u_B, u_C)$
versus
A
B
C

Non-classical

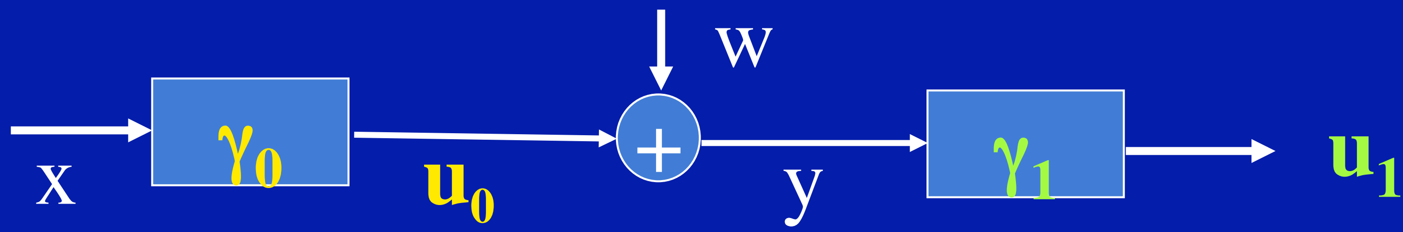


$x \sim N(0, \sigma_x^2)$ $w \sim N(0, \sigma_w^2)$

$J(\gamma_0, \gamma_1) = E[\, Q(x, u_0, u_1) \mid \gamma_0, \gamma_1 \,]$

$J^* = \min \min J(\gamma_0, \gamma_1)$

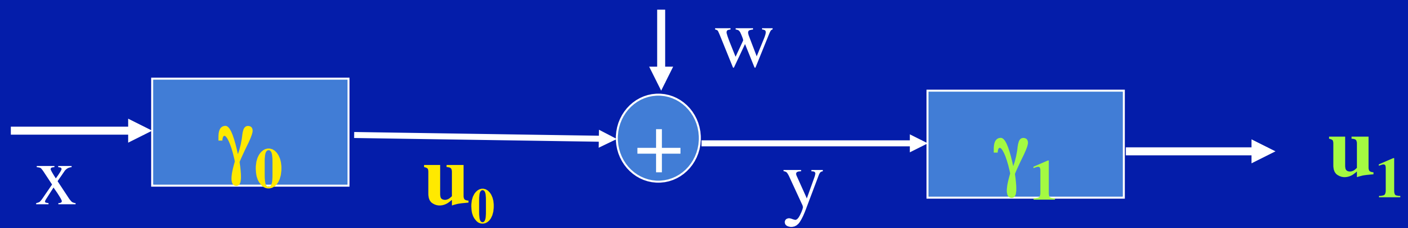
Witsenhausen (1968)



$$Q_W(x, u_0, u_1) = k_0 (u_0 - x)^2 + (u_0 - u_1)^2$$

optimal control law exists, but its structure is not known

Witsenhausen (1968)



$$Q_W(x, u_0, u_1) = k_0\,(u_0 - x)^2 + (u_0 - u_1)^2$$

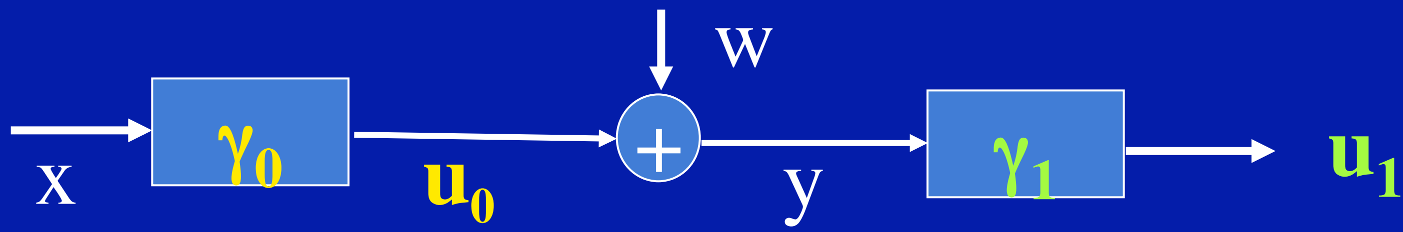
A control law that beats the best linear one:

$$u_0 = \gamma_0(x) = \varepsilon\,\mathrm{sgn}\,(x) + \lambda\,x$$

$$u_1 = \gamma_1(y) = E[\varepsilon\,\mathrm{sgn}\,(x) + \lambda\,x \mid y]$$

optimize wrt ε and λ

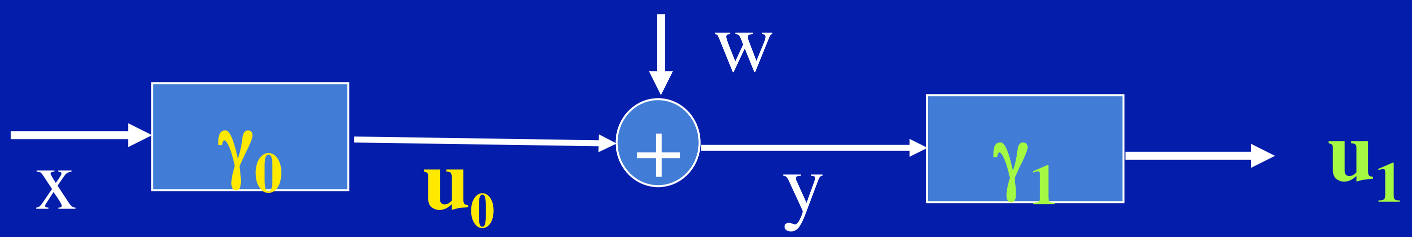
Gaussian Test Channel



$$Q_{TC}(x, u_0, u_1) = k_0 (u_0)^2 + (u_1 - x)^2$$

optimal control law (encoder/decoder) exists, and is linear

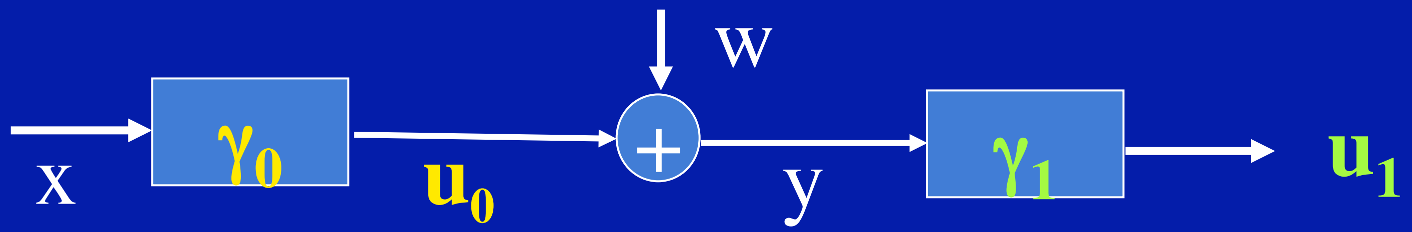
Generalized Gaussian Test Channel



$$Q_{GTC}(x, u_0, u_1) = k_0 (u_0)^2 + (u_1 - x)^2 + b_0 u_0 x$$

⇒ optimal control law (encoder/decoder) exists, and is linear

Generalized Gaussian Test Channel

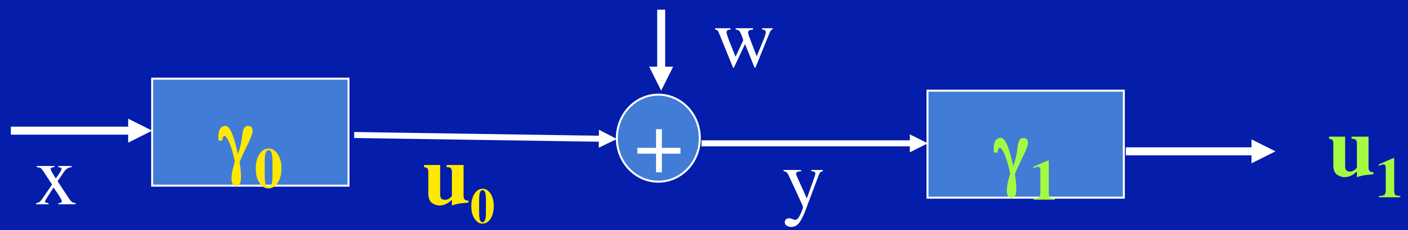


$$Q(x, u_0, u_1) = k_0 (u_0)^2 + (u_1 - x)^2 + b_0 u_0 x$$

$$E[Q]=F(\gamma_0, \gamma_1) \geq k_0 \alpha + \beta + \inf_\gamma b_0 E[\gamma_0(x)x]$$

$$\geq k_0 \alpha + \beta - \mathrm{sgn}(b_0)\, \sigma_x \sqrt{\alpha}$$

Generalized Gaussian Test Channel



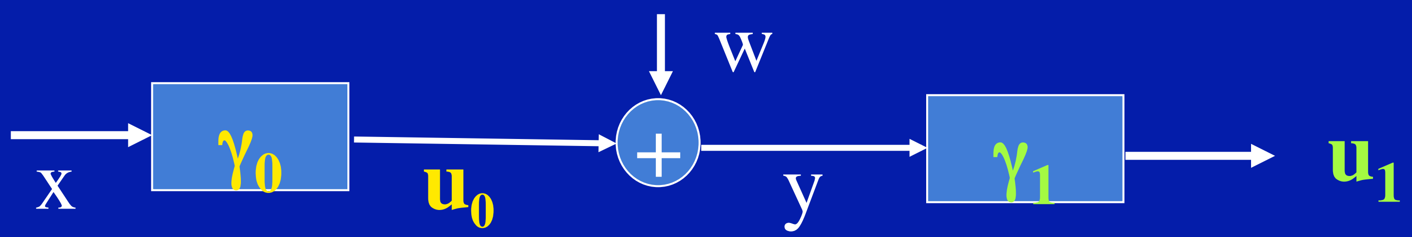
$$Q(x, u_0, u_1) = k_0 (u_0)^2 + (u_1 - x)^2 + b_0 u_0 x$$

$$E[Q]=F(\gamma_0, \gamma_1) \geq k_0 \alpha + \beta + \inf_\gamma b_0 E[\gamma_0(x)x]$$

$$\geq k_0 \alpha + \beta - \mathrm{sgn}(b_0)\, \sigma_x \sqrt{\alpha}$$

DPT: $I(U_0;Y) \geq I(X;U_1)$

Generalized Gaussian Test Channel



$$Q(x, u_0, u_1) = k_0 (u_0)^2 + (u_1 - x)^2 + b_0 u_0 x$$

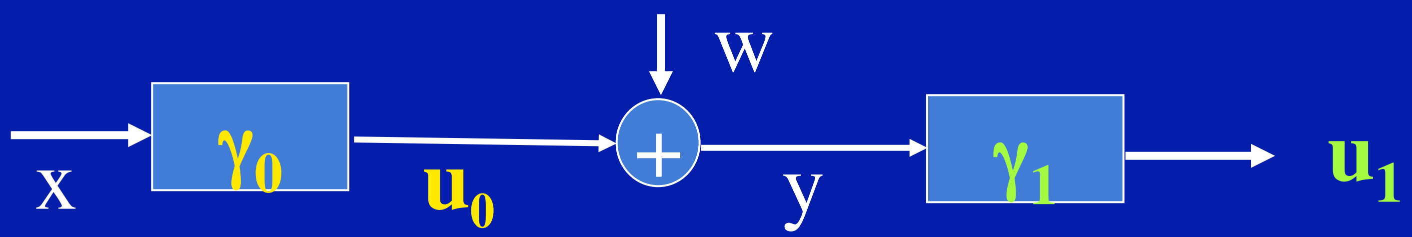
$$E[Q]=F(\gamma_0, \gamma_1) \geq k_0 \alpha + \beta + \inf_\gamma b_0 E[\gamma_0(x)x]$$

$$\geq k_0 \alpha + \beta - \mathrm{sgn}(b_0)\, \sigma_x \sqrt{\alpha}$$

$$(1/2)\log(1+(\alpha/\sigma_w^2)) \geq I(U_0;Y) \geq I(X;U_1) \geq (1/2)\log(\sigma_x^2/\beta)$$

$C(\alpha)$ $R(\beta)$

Generalized Gaussian Test Channel



$$Q(x, u_0, u_1) = k_0 (u_0)^2 + (u_1 - x)^2 + b_0 u_0 x$$

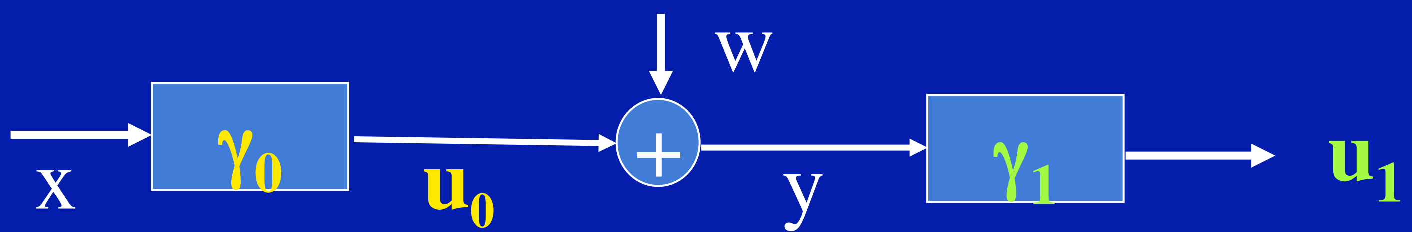
$$E[Q]=F(\gamma_0, \gamma_1) \geq k_0 \alpha + \beta + \inf_\gamma b_0 E[\gamma_0(x)x]$$

$$\geq k_0 \alpha + \beta - \mathrm{sgn}(b_0) \sigma_x \sqrt{\alpha}$$

$$(1/2)\log (1+(\alpha/ \sigma_w^2)) \geq I(U_0;Y) \geq I(X;U_1) \geq (1/2)\log (\sigma_x^2/ \beta)$$

==> $\beta \geq \sigma_w^2 \sigma_x^2 / (\sigma_w^2 + \alpha)$

Generalized Gaussian Test Channel



$$Q(x, u_0, u_1) = k_0 (u_0)^2 + (u_1 - x)^2 + b_0 u_0 x$$

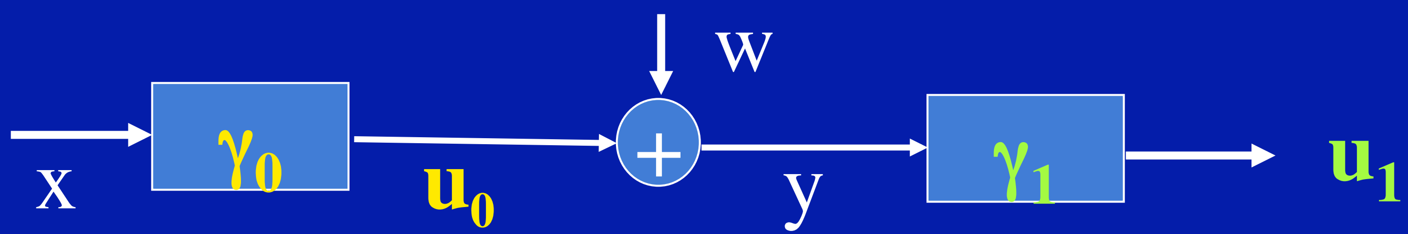
$$E[Q]=F(\gamma_0, \gamma_1) \geq k_0 \alpha + \beta + \inf_\gamma b_0 E[\gamma_0(x)x]$$

$$\geq k_0 \alpha + \beta - \mathrm{sgn}(b_0)\, \sigma_x \sqrt{\alpha}$$

==> $\beta \geq \sigma_w^2 \sigma_x^2 / (\sigma_w^2 + \alpha)$

Inequality is tight with $\gamma_0(x) = -\mathrm{sgn}(b_0)(\sqrt{\alpha} / \sigma_x)\, x$

Generalized Gaussian Test Channel



$$Q(x, u_0, u_1) = k_0 (u_0)^2 + (u_1 - x)^2 + b_0 u_0 x$$

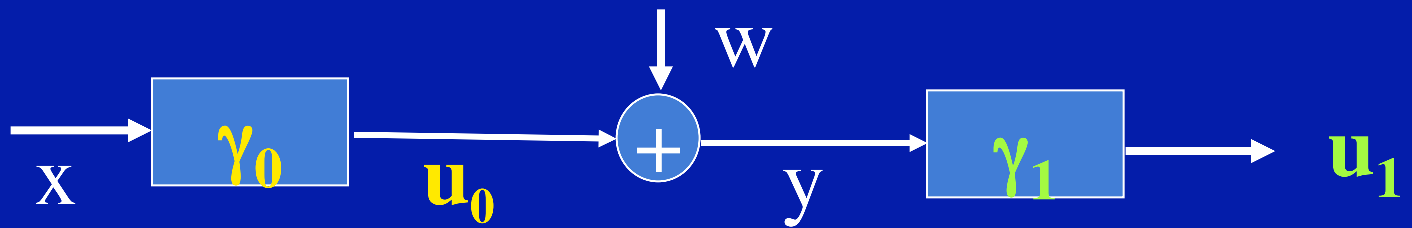
$$E[Q] = F(\gamma_0, \gamma_1) \geq k_0 \alpha + \beta - |b_0| \sigma_x \sqrt{\alpha}$$

$$\geq k_0 \alpha + \sigma_w^2 \sigma_x^2 / (\sigma_w^2 + \alpha) - |b_0| \sigma_x \sqrt{\alpha}$$

Obtain the α that minimizes the bound --> α^*

Then, $\gamma_0^*(x) = -\mathrm{sgn}(b_0)(\sqrt{\alpha^*} / \sigma_x)\, x$, $\gamma_1^*(y) = E[x|y]$

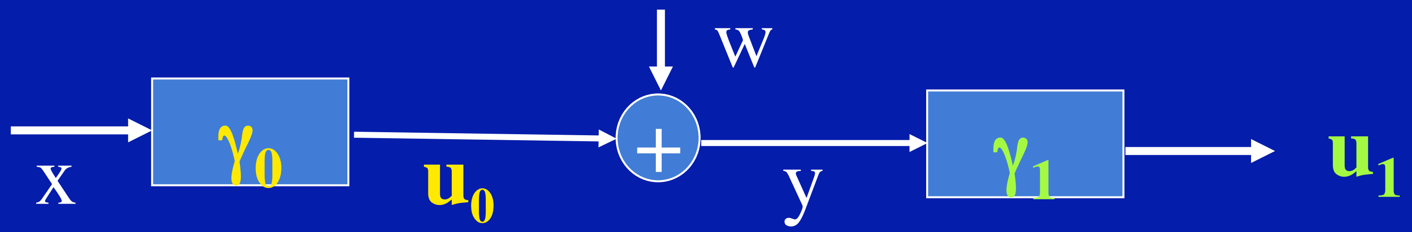
Generalized Gaussian Test Channel



$$Q(x, u_0, u_1) = k_0 (u_0)^2 + (u_1 - x)^2 + b_0 u_0 x$$

One of the *few* instances when static/causal coding (and linear in this case) leads to attainment of equality in $C(\alpha) \geq R(\beta)$

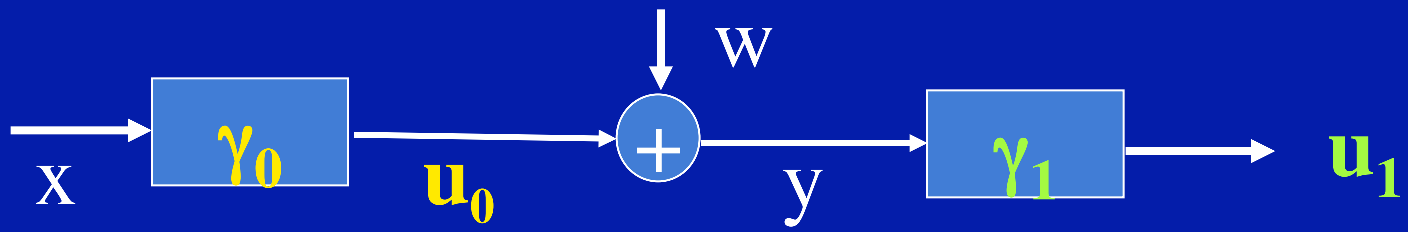
Revisit: Witsenhausen (1968)



$$Q(x, u_0, u_1) = k_0 (u_0 - x)^2 + (u_0 - u_1)^2$$

Because of the product term $u_0 u_1$ the preceding analysis does not apply here

However, with Conflicting Objectives

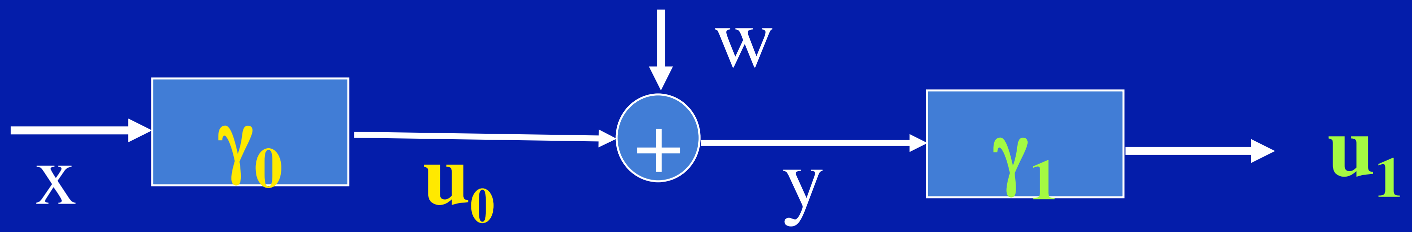


$$Q_G(x, u_0, u_1) = -\, k_0\,(u_0 - x)^2 + (u_0 - u_1)^2$$

$$J_* = \min_{\gamma_1}\ \max_{\gamma_0} J(\gamma_0\ , \gamma_1)$$

Unique saddle-point solution, control laws are linear

However, with Conflicting Objectives



$$Q_G(x, u_0, u_1) = - k_0 (u_0 - x)^2 + (u_0 - u_1)^2$$

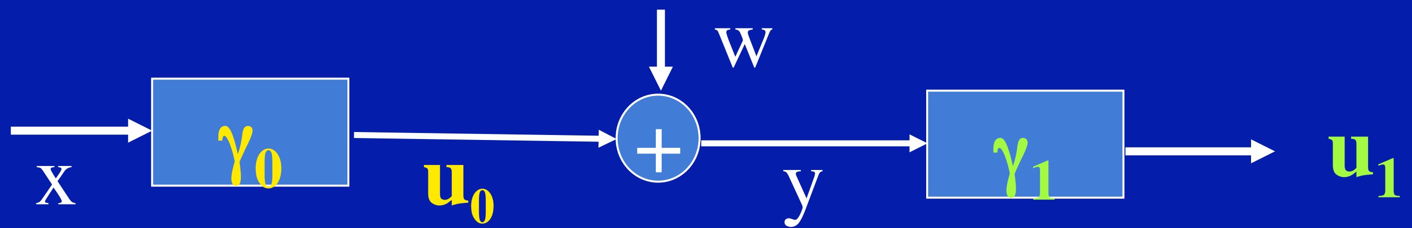
$\gamma_0^*(x) = - [k_0 / (k_0 - (\lambda^* - 1)^2)]x, \; \gamma_1^*(y) = \lambda^* y$

where λ^* **uniquely** solves the polynomial eq

$f(\lambda) = (\sigma_w^2 / \sigma_x^2)\, \lambda[k_0 - (\lambda - 1)^2]^2 - k_0^2(1-\lambda) = 0$

in the open interval $(\max(0, 1-\sqrt{k_0}), 1)$

Recap



$Q_W = \ k_0 (u_0 - x)^2 + (u_0 - u_1)^2$ conflicting roles

$Q_G = - k_0 (u_0 - x)^2 + (u_0 - u_1)^2$ aligned roles

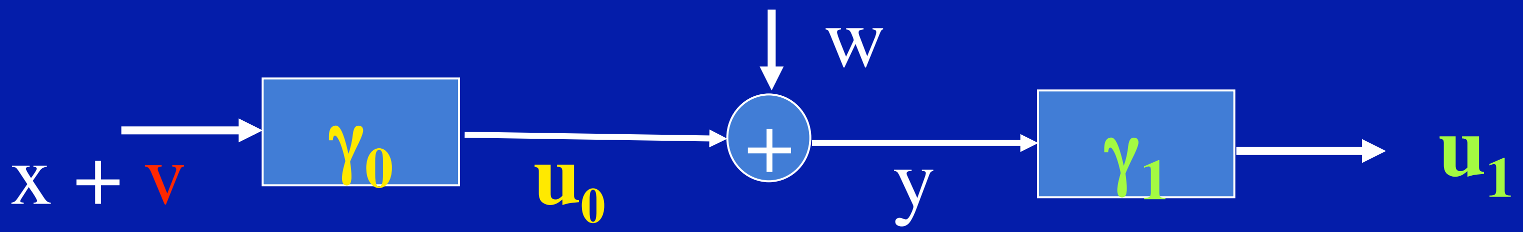
$Q_{TC} = k_0 (u_0)^2 + (u_1 - x)^2$ aligned roles

⇒ Not only the information structure but also the cost function is a determining factor

Extensions of the Paradigm

- Noise corrupted access to initial state
- Vector-valued variables
- Stochastic LQG teams
- Non-cooperative games

Noise Corrupted IS

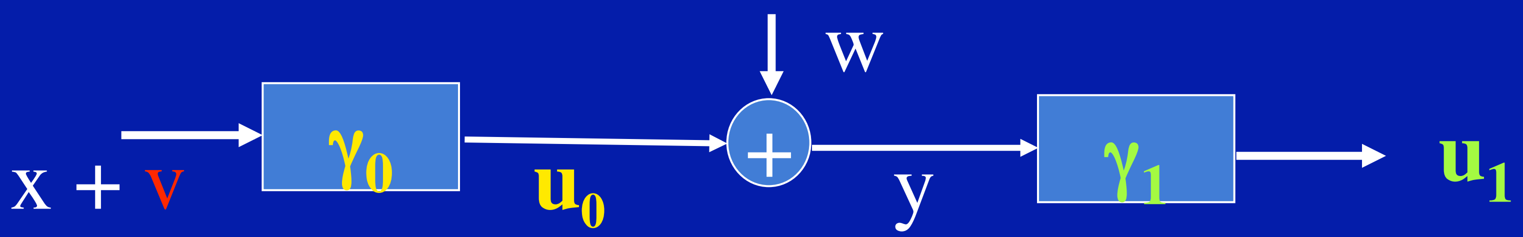


$$x \sim N(0, \sigma_x^2), \quad w \sim N(0, \sigma_w^2), \quad v \sim N(0, \sigma_v^2)$$

$$J(\gamma_0\ , \gamma_1) = E\ [\ Q(x, u_0, u_1) \mid \gamma_0\ , \gamma_1\ ]$$

➔ Similar structural results

Noise Corrupted IS



GTC: for some unique positive α^*

$\gamma_0^*(z) = \alpha^* \mathbf{z}$, $\gamma_1^*(\mathbf{y}) = E[\mathbf{x}|\mathbf{y}]$; $\mathbf{z} := \mathbf{x}+\mathbf{v}$

ZSSG: for some λ^*, root of a 5th-order polynomial

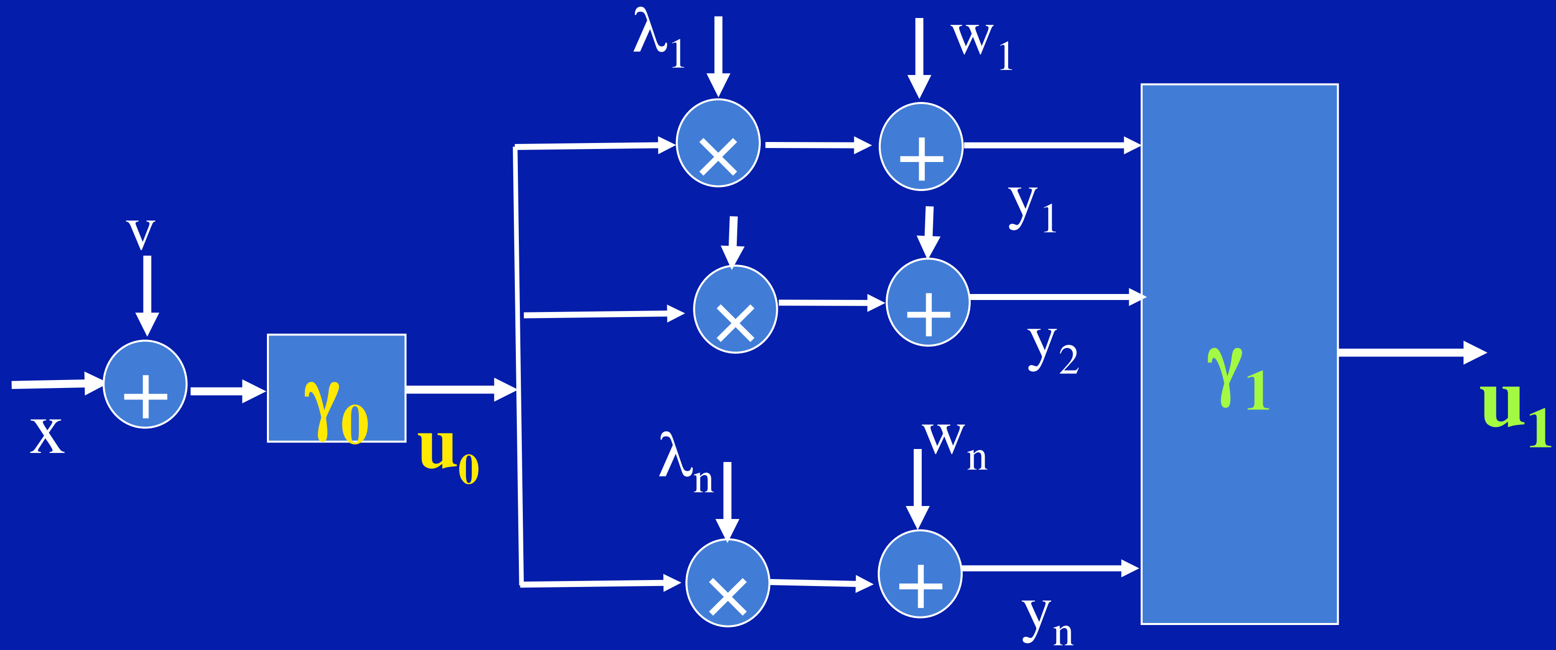
$\gamma_0^*(\mathbf{z}) = -[k_0 / (k_0 - (\lambda^* - 1)^2)][\sigma_x^2/(\sigma_x^2 + \sigma_v^2)]\mathbf{z}$

$\gamma_1^*(\mathbf{y}) = \lambda^* \mathbf{y}$

Vector-Valued Variables

- Additional difficulties even for GTC, unless decision variables are scalar but channels are vector-valued (next)
- ZSSG is still tractable, and unique SP solution is linear

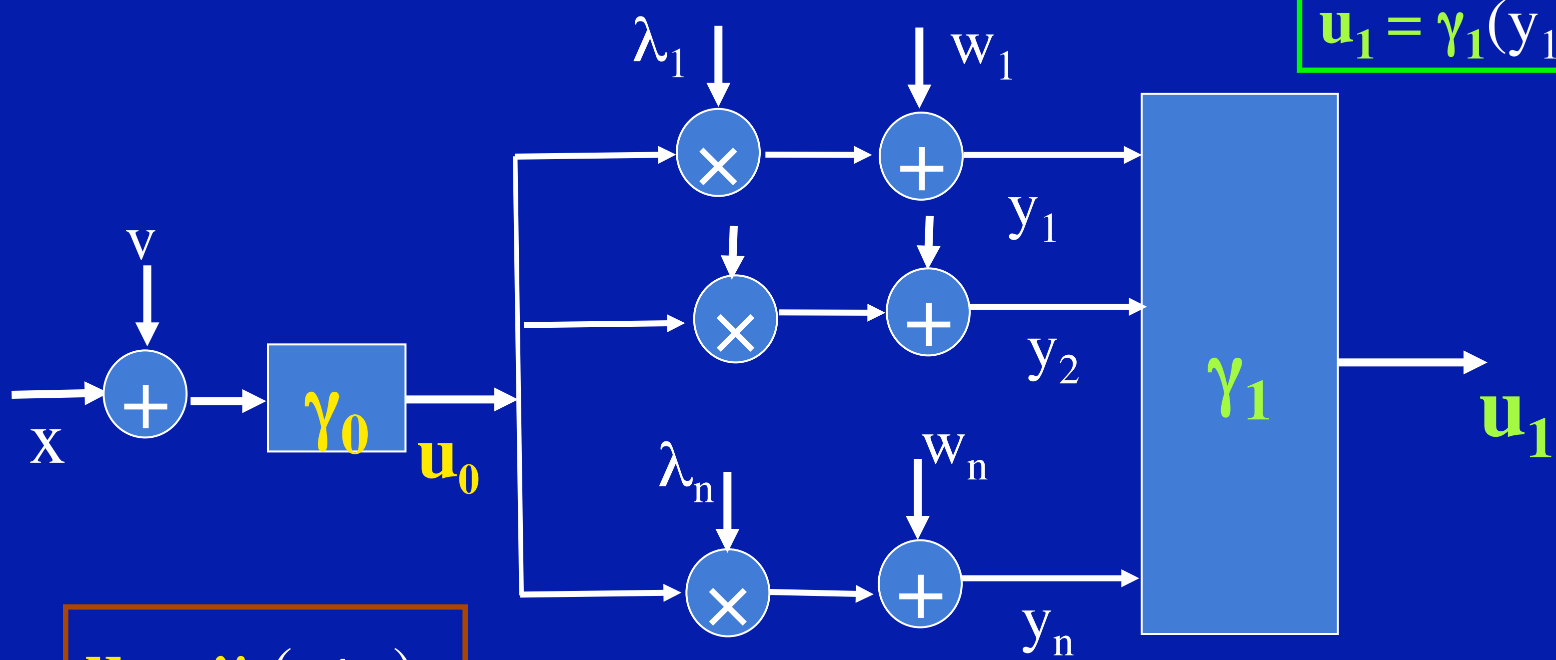
A multi-channel extension to GTC



λ_i's are nonzero constants (gains);
x, v, w_i's are independent, Gaussian random variables

A multi-channel extension to GTC

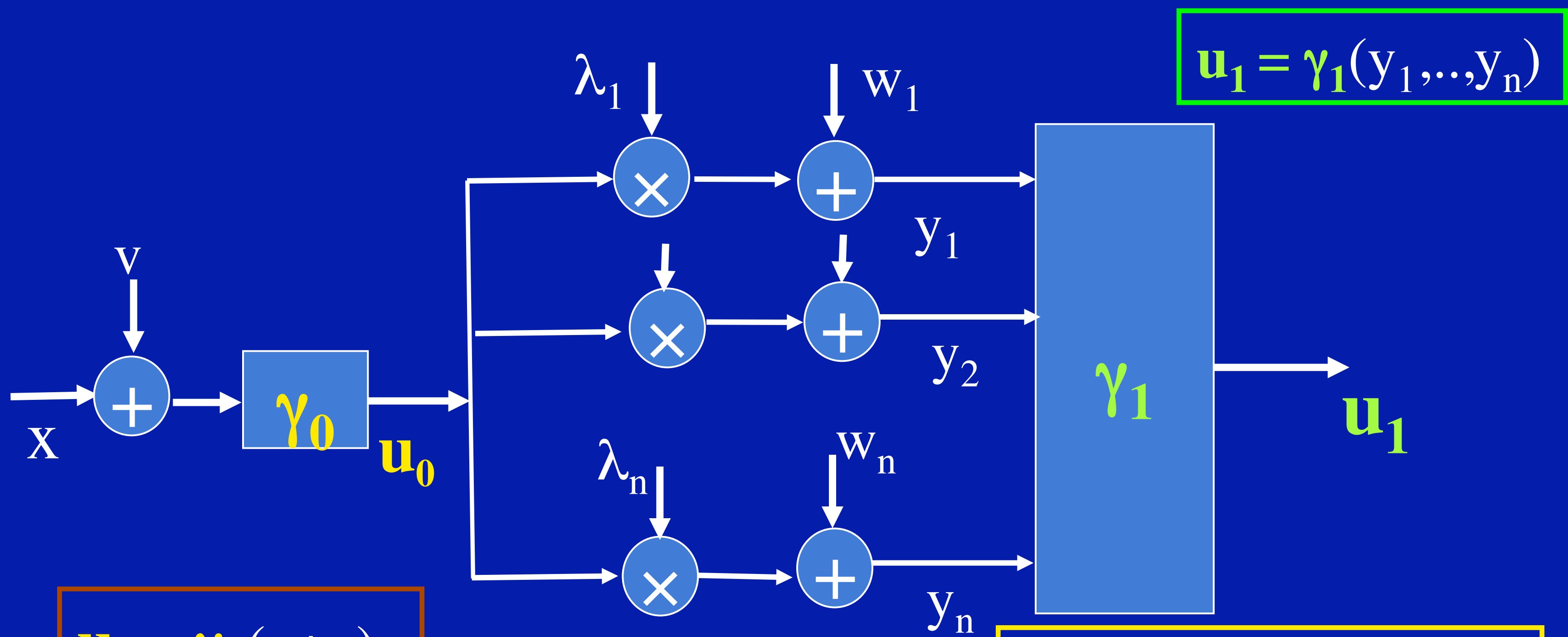
$u_1 = \gamma_1(y_1,..,y_n)$



$u_0 = \gamma_0(x+v)$

$y_i = \lambda_i \gamma_0(x+v) + w_i$

A multi-channel extension to GTC



$u_0 = \gamma_0(x+v)$

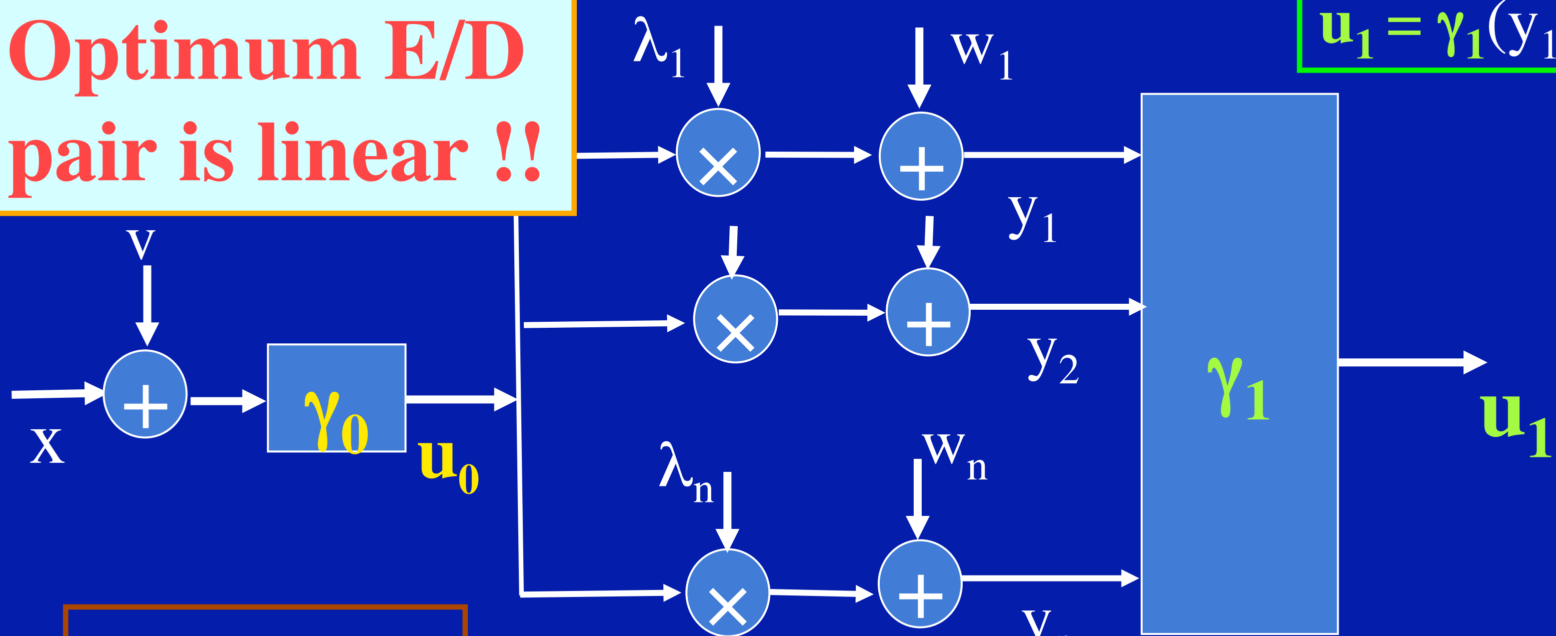
$y_i = \lambda_i \gamma_0(x+v) + w_i$

$$Q_{GTC} = k_0 (u_0)^2 + (u_1 - x)^2 + b_0 u_0 x$$

A multi-channel extension to GTC

Optimum E/D pair is linear !!

$u_1 = \gamma_1(y_1,..,y_n)$



$u_0 = \gamma_0(x+v)$

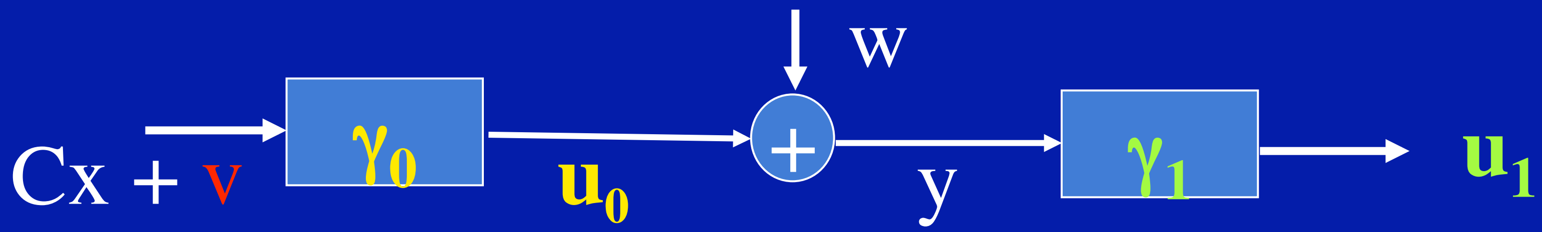
$y_i = \lambda_i \gamma_0(x+v) + w_i$

$$Q_{GTC} = k_0 (u_0)^2 + (u_1 - x)^2 + b_0 u_0 x$$

Stochastic LQG Teams

- To make tractable, one needs a *forward* channel that informs agents at the front end on garbled decentralized information received at the back end → *quasi-classical*
- $\gamma_{0i}(z_i)$ at back end, i=1, …, n
- $\gamma_{1i}(y_i, z)$ at front end i=1, …, n
- For quadratic teams invoke *Radner (62) and extensions*

Vector-Valued Decision Variables (Decentralized)



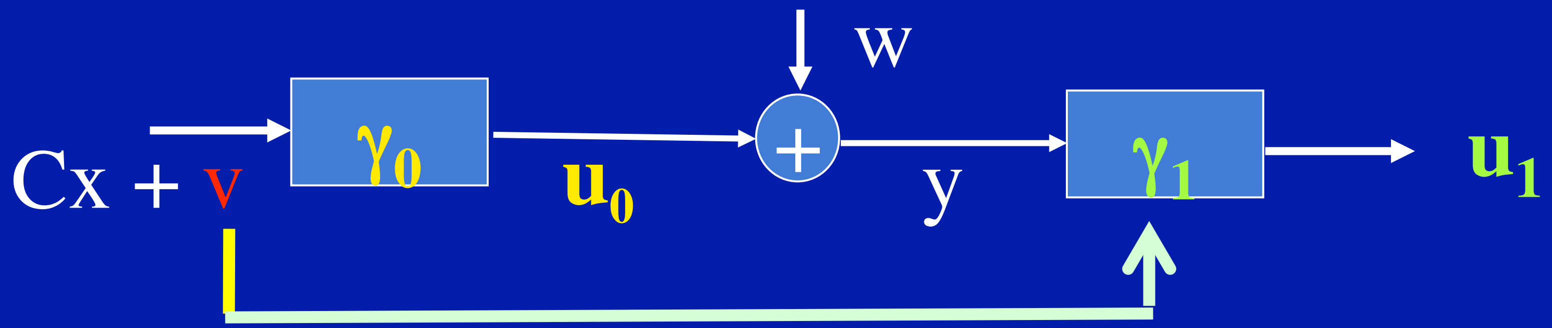
$u_{0i} = \gamma_{0i}(z_i)$, $z_i = C_i x + v_i$, $i=1, \dots, n$

$u_{1i} = \gamma_{1i}(z, y_i)$, $y_i = D_i u_0 + w_i$, $i=1, \dots, n$

$z = (z_1, \dots , z_n)$; w correlated with x

$J(\gamma_0 , \gamma_1) = E [Q(x, u_0, u_1) \mid \gamma_0 , \gamma_1]$

Vector-Valued Decision Variables (Decentralized)

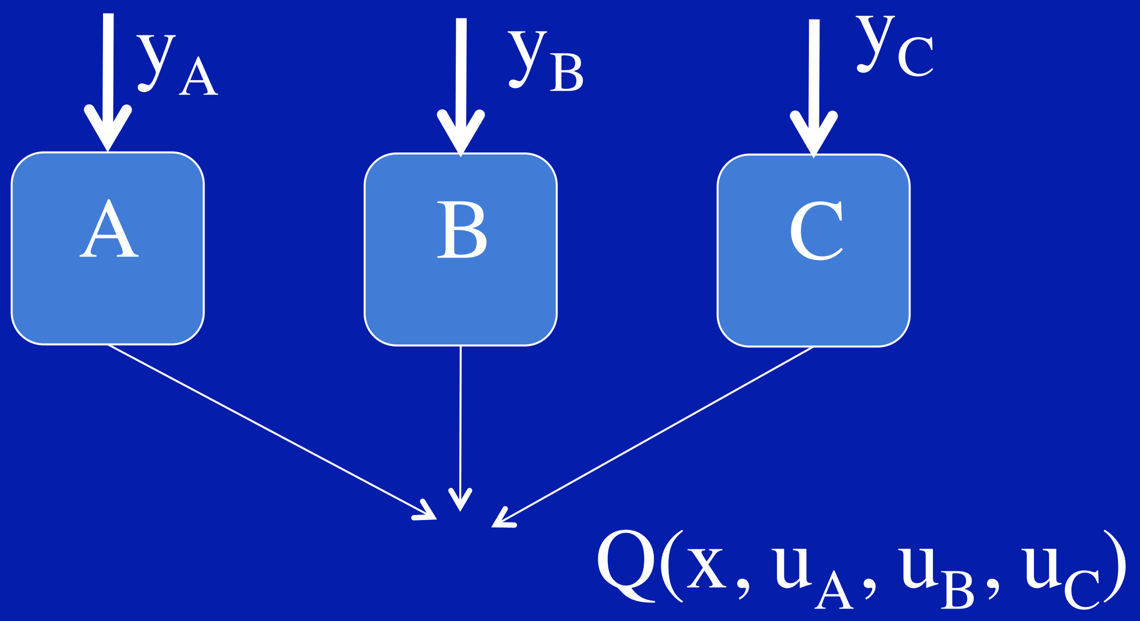


$u_{0i} = \gamma_{0i}(z_i)$, $z_i = C_i x + v_i$, $i=1, \ldots, n$

$u_{1i} = \gamma_{1i}(z, y_i)$, $y_i = D_i u_0 + w_i$, $i=1, \ldots, n$

$z = (z_1, \ldots , z_n)$; w correlated with x

$J(\gamma_0 , \gamma_1) = E [Q(x, u_0, u_1) \mid \gamma_0 , \gamma_1]$



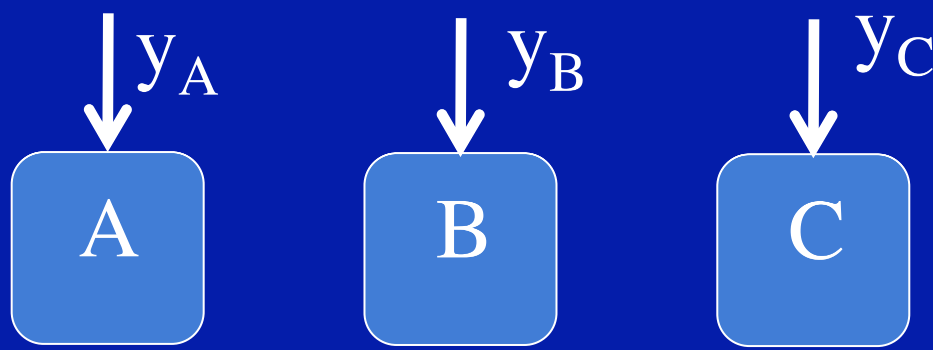
Radner (62): y's jointly Gaussian distributed, Q strictly (jointly) convex

➔ unique team optimal solution

Stochastic Nash Games

- Again one needs a *forward* channel that informs agents at the front end on garbled decentralized information received at the back end (but not actions) → *quasi-classical*
- $\gamma_{0i}(z_i)$ at back end, i=1, …, n
- $\gamma_{1i}(y_i, z)$ at front end i=1, …, n
- For quadratic games use *TB (74, 75, 78) as extension of Radner (62)*

Stochastic Nash Games



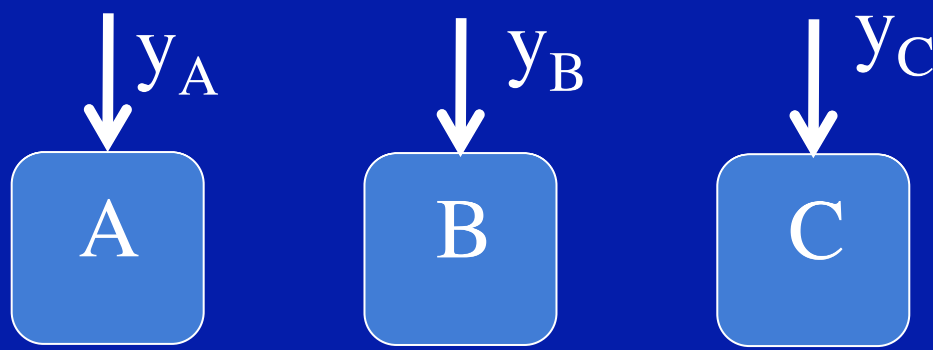
$$Q_i(x, u_A, u_B, u_C), \quad i = A, B, C$$

Nash eqm: $(\Upsilon_A, \Upsilon_B, \Upsilon_C)$

Υ_A minimizes $J_A(\Upsilon_A, \Upsilon_B, \Upsilon_C)$;

likewise for B, C

Stochastic Nash Games



$Q_i(x, u_A, u_B, u_C)$, $i = A, B, C$

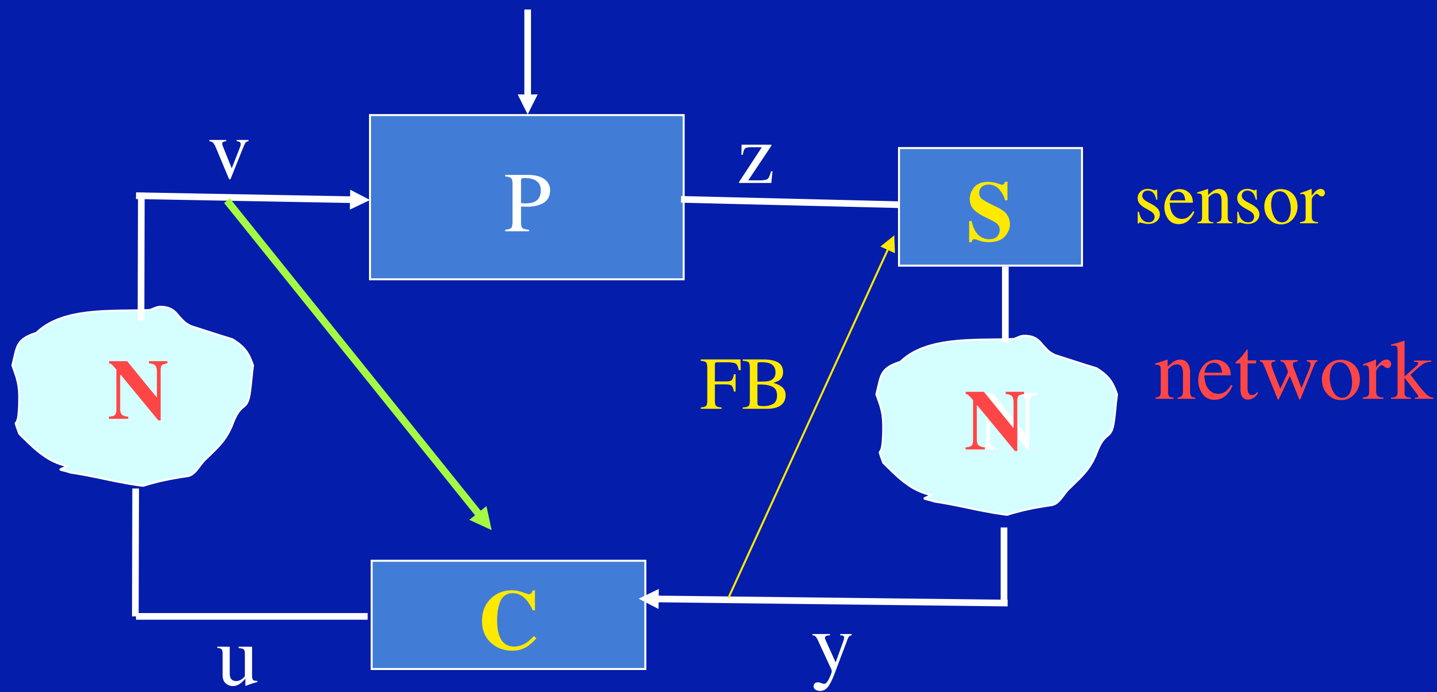
TB (74, 75, 78): y's jointly Gaussian distributed, Q_i strictly convex + technical condition

→ unique Nash eqm solution; linear

Extension to Multi-Stage Scenarios

Dynamic Systems

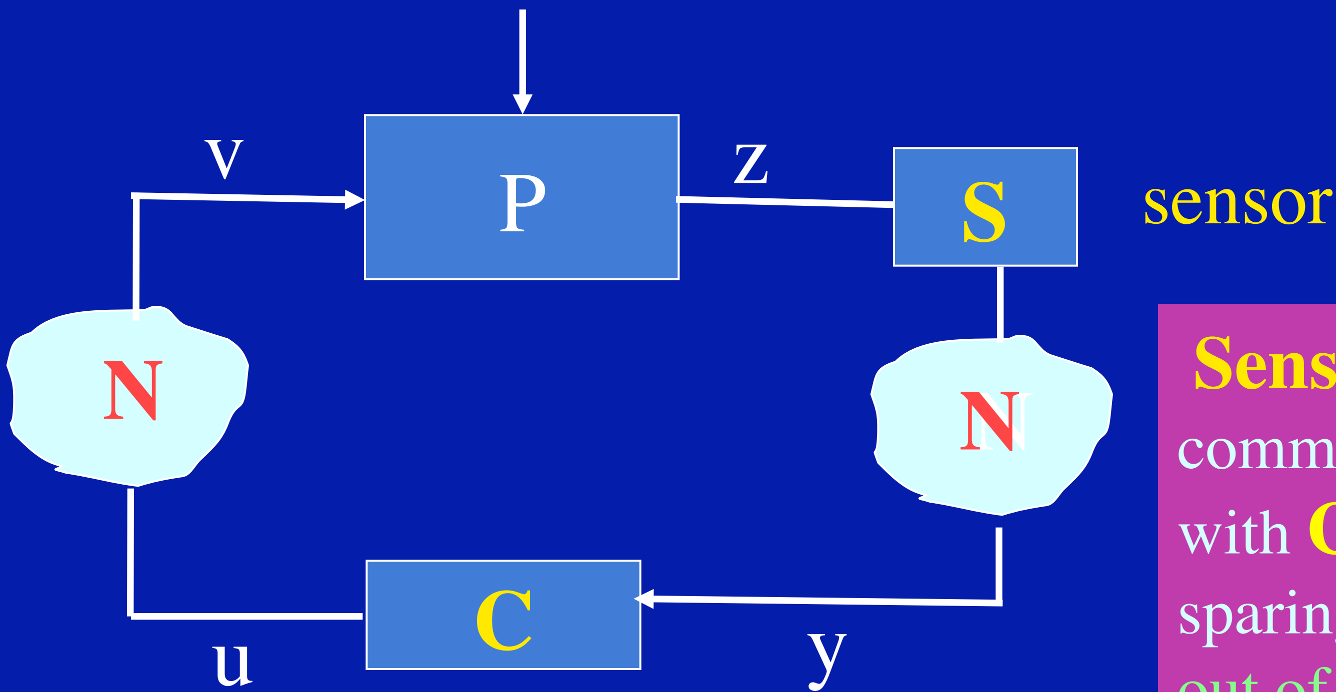
Remote Control Paradigm



PI (S, C) → optimize

Non-classical information!

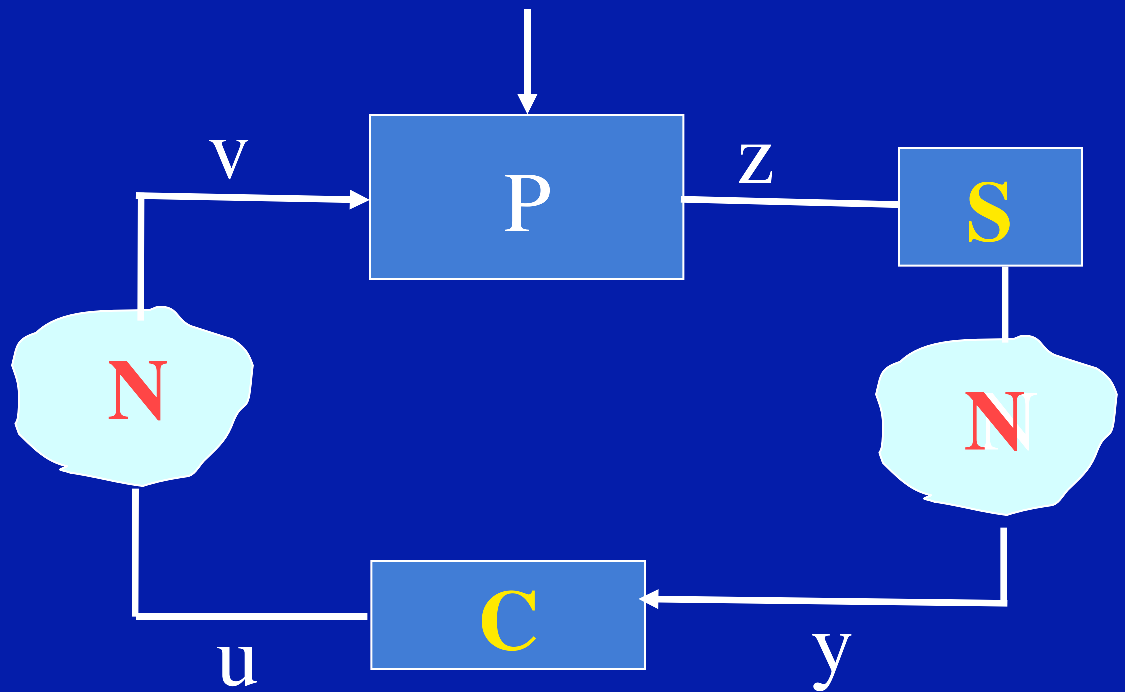
Limited Usage



Sensor communicates with **Control** sparingly: M out of N times

PI → optimize

Limited Usage



sensor

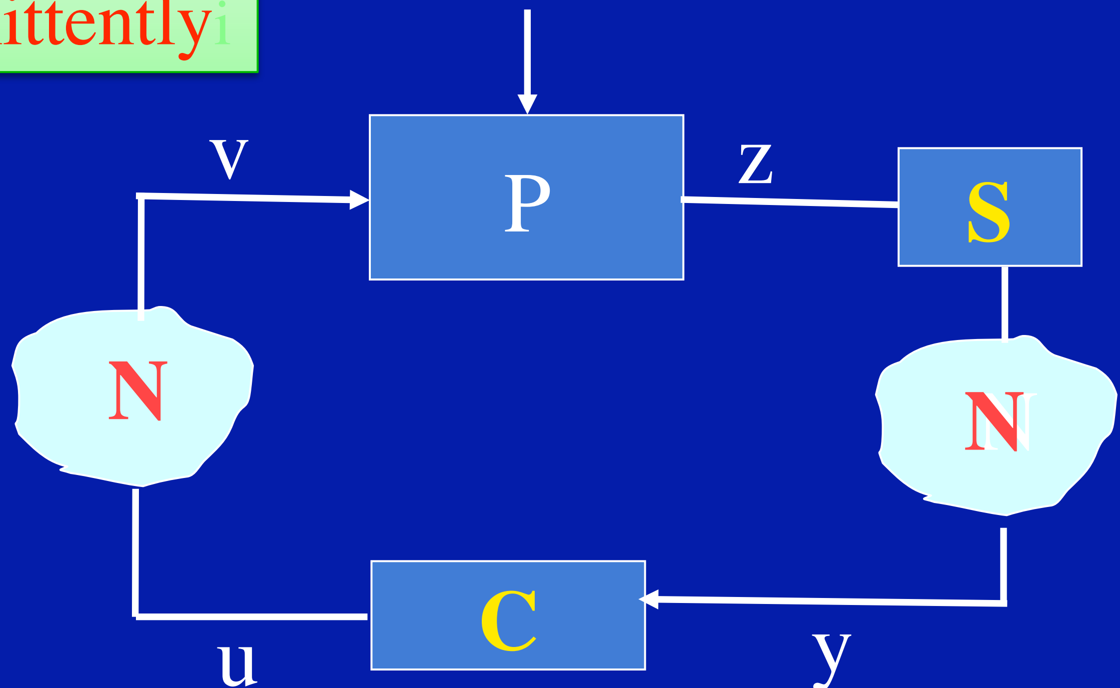
Sensor communicates with **Control** sparingly: M out of N times

Controller communicates with **Plant** intermittently

PI → optimize

Limited Usage

Jammer
disrupts
intermittentlyi



sensor

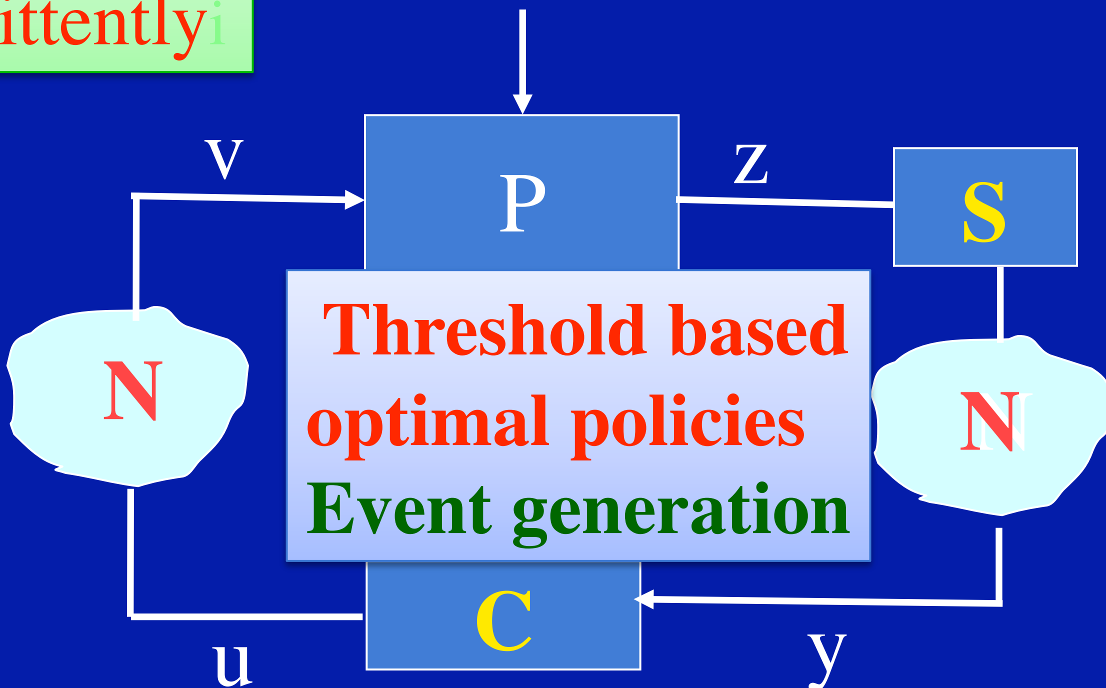
Sensor
communicates
with **Control**
sparingly: M
out of N times

Controller
communicates
with **Plant**
intermittently

PI → optimize

Limited Usage

Jammer disrupts intermittently



sensor

Threshold based optimal policies

Event generation

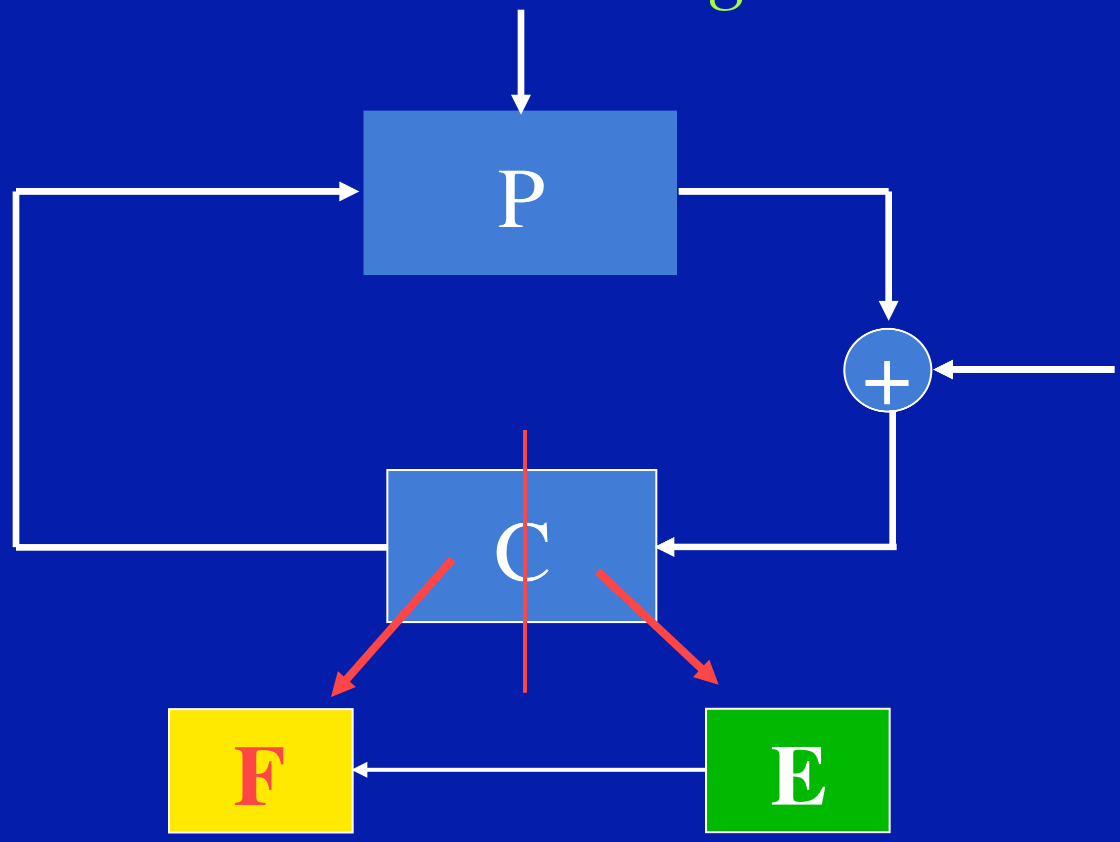
Sensor communicates with **Control** sparingly: M out of N times

Controller communicates with **Plant** intermittently

PI → optimize

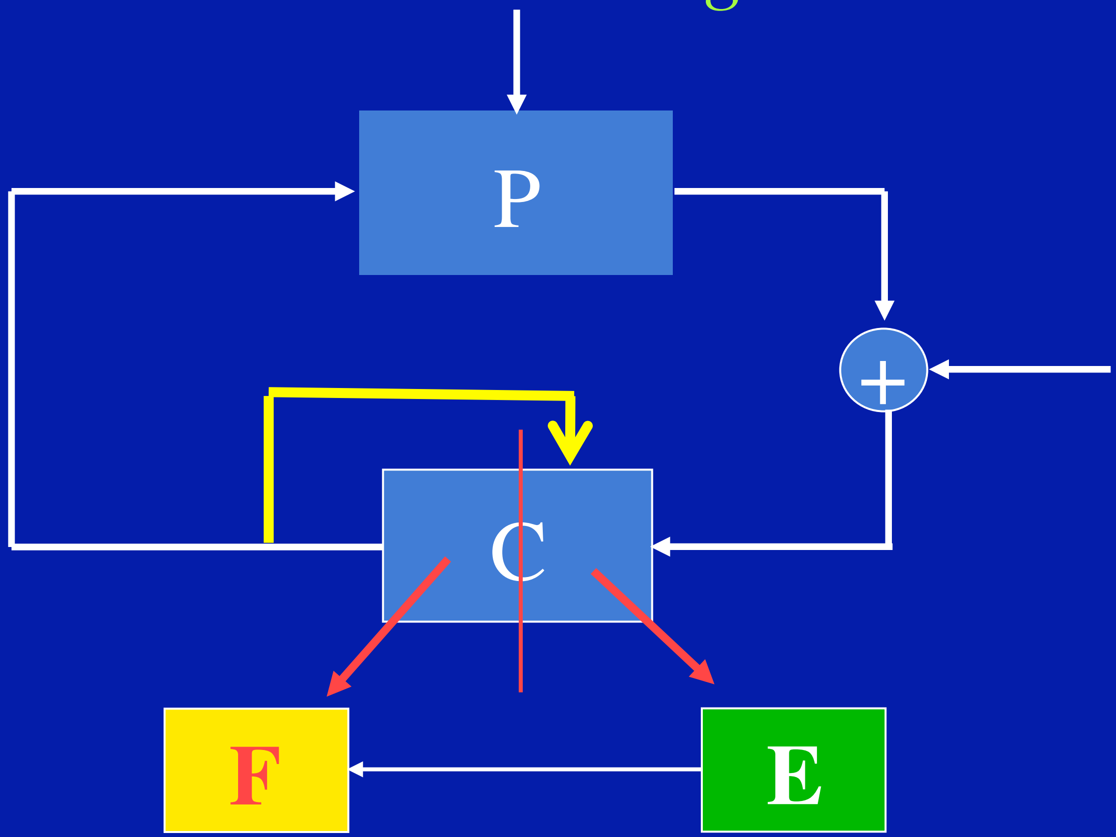
Back to Separation / Neutrality

Does it hold in games?



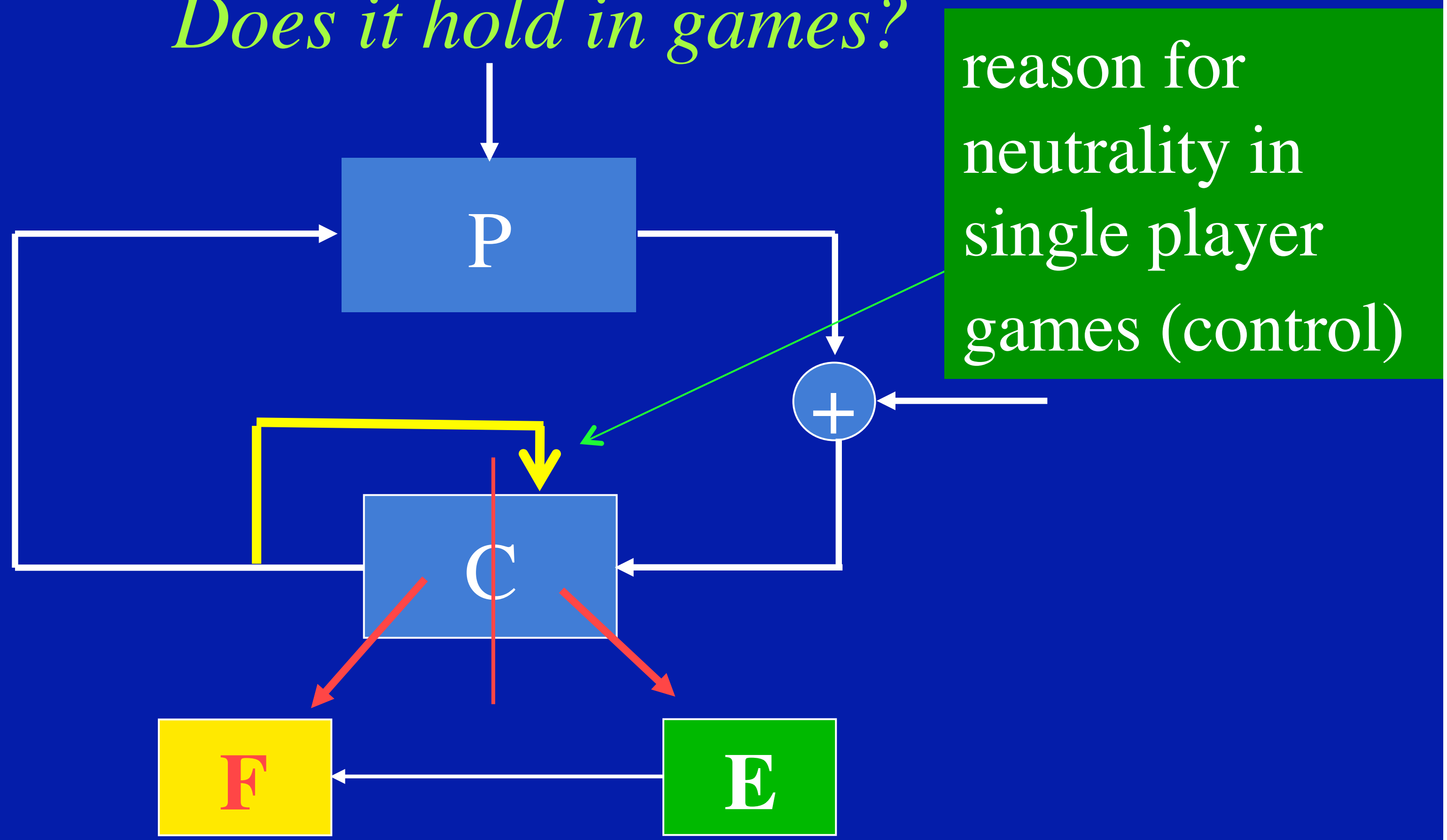
Back to Separation / Neutrality

Does it hold in games?



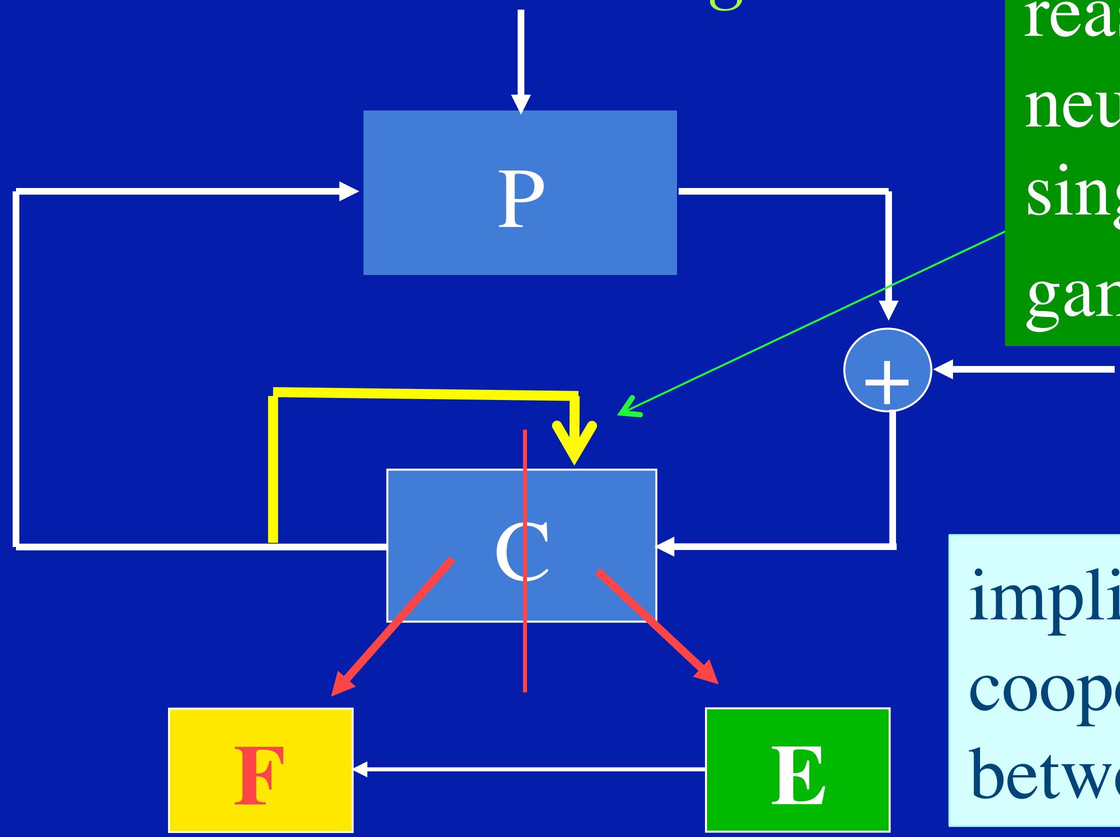
Back to Separation / Neutrality

Does it hold in games?



Back to Separation / Neutrality

Does it hold in games?



reason for neutrality in single player games (control)

implicit cooperation between F & E

ZSSDG with common measurements

$dx_t = (Ax_t + Bu_t + Dv_t)\,dt + F\,d\xi_t\,, \quad t \geq 0$

$dy_t = Hx_t\,dt + G\,dw_t\,, \quad y_0 = 0$ (common measurement)

$u_t = \gamma_t\,(y_0^{\,t}) \qquad\qquad v_t = \mu_t\,(y_0^{\,t})$

$PI = E\{ \int_0^{tf} [\, |x_t|_Q^{\,2} + |u_t|^2 - |v_t|^2 \,]\, dt + |x_{tf}|_{Qf}^{\,2} \}$

$\min_\gamma \max_\mu \; J(\gamma, \mu) \;\rightarrow\; (\gamma^*, \mu^*)$ say SP

ZSSDG with common measurements

$dx_t = (Ax_t + Bu_t + Dv_t)\, dt + F\, d\xi_t\,, \;\; t \geq 0$

$dy_t = Hx_t\, dt + G\, dw_t\,, \;\; y_0 = 0$ (common measurement)

$u_t = \gamma_t\,(y_0^{\,t})$ $\qquad v_t = \mu_t\,(y_0^{\,t})$

$PI = E\{ \int_0^{t_f} [\, |x_t|_Q^{\,2} + |u_t|^2 - |v_t|^2 \,]\, dt + |x_{tf}|_{Qf}^{\,2} \}$

$\min_\gamma \max_\mu \; J(\gamma, \mu) \;\rightarrow\; (\gamma^*, \mu^*)$

Does certainty equivalence hold?

-- can SP policies from deterministic game be used?

ZSSDG with common measurements

$dx_t = (Ax_t + Bu_t + Dv_t)\, dt + F\, d\xi_t\,, \;\; t \geq 0$

$dy_t = Hx_t\, dt + G\, dw_t\,, \;\; y_0 = 0$ (common measurement)

$u_t = \gamma_t\,(y_0^{\,t})$ $\qquad$ $v_t = \mu_t\,(y_0^{\,t})$

$PI = E\{ \int_0^{t_f} [\, |x_t|_Q^{\,2} + |u_t|^2 - |v_t|^2 \,]\, dt + |x_{tf}|_{Qf}^{\,2} \}$

$\min_\gamma \max_\mu \; J(\gamma, \mu) \;\rightarrow\; (\gamma^*, \mu^*)$

Does certainty equivalence hold? *Qualified* NO

Building a common filter with u, v requires cooperation

ZSSDG with common measurements

$dx_t = (Ax_t + Bu_t + Dv_t)\,dt + F\,d\xi_t\,,\ \ t \geq 0$

$dy_t = Hx_t\,dt + G\,dw_t\,,\ \ y_0 = 0$ (common measurement)

$u_t = \gamma_t\,(y_0^{\,t})$ $\qquad v_t = \mu_t\,(y_0^{\,t})$

$PI = E\{ \int_0^{t_f} [\ |x_t|_Q^2 + |u_t|^2 - |v_t|^2\ ]\,dt + |x_{tf}|_{Qf}^2 \}$

$\min_\gamma \max_\mu\ J(\gamma, \mu)$ ➔ (γ^*, μ^*)

Still, there exists a common compensator, and restricted CE/separation holds -- but not complete

NZSSDG with common measurements

$dx_t = (Ax_t + Bu_t + Dv_t)\, dt + F\, d\xi_t\,,\ \ t \geq 0$

$dy_t = Hx_t\, dt + G\, dw_t\,,\quad y_0 = 0$ (common measurement)

$u_t = \gamma_t\,(y_0^{\,t})$ $\qquad\qquad v_t = \mu_t\,(y_0^{\,t})$

$PIi = E\{ \int_0^{t_f} [\ |x_t|_{Qi}^{\;2} + |u_t|_{Ri}^{\;2} + |v_t|_{Mi}^{\;2}\ ]\, dt + |x_{tf}|_{Qfi}^{\;2} \}$

➔ $J_i(\gamma, \mu)$ ➔ Nash eqm (γ^*, μ^*)

NZSSDG with common measurements

$dx_t = (Ax_t + Bu_t + Dv_t)\,dt + F\,d\xi_t\,,\ \ t \geq 0$

$dy_t = Hx_t\,dt + G\,dw_t\,,\ \ y_0 = 0$ (common measurement)

$u_t = \gamma_t\,(y_0^{\,t})$ $\qquad$ $v_t = \mu_t\,(y_0^{\,t})$

$PIi = E\{ \int_0^{t_f} [\ |x_t|_{Qi}^2 + |u_t|_{Ri}^2 + |v_t|_{Mi}^2\ ]\,dt + |x_{tf}|_{Qfi}^2 \}$

→ $J_i(\gamma, \mu)$ → Nash eqm (γ^*, μ^*)

CE/separation does not hold -- NE of deterministic NZSDG cannot be used; not neutral

Recap

- No general theory/approach to non-neutrality
- Not all problems with non-classical information are intractable
- It is not only the information structure but also the structure of the performance index that plays an important role in tractability vs intractability
- With battery limitations and energy conservation in multi agent applications, further research on problems with non-classical information is needed

THANKS !