Sensing, Coordination and Control in Adversarial Environments with Limited Actions

TAMER BAŞAR

ECE, CAS, CSL, ITI, and MechSE, UIUC

basar1@illinois.edu

30th Chinese Control Conference (CCC'11)

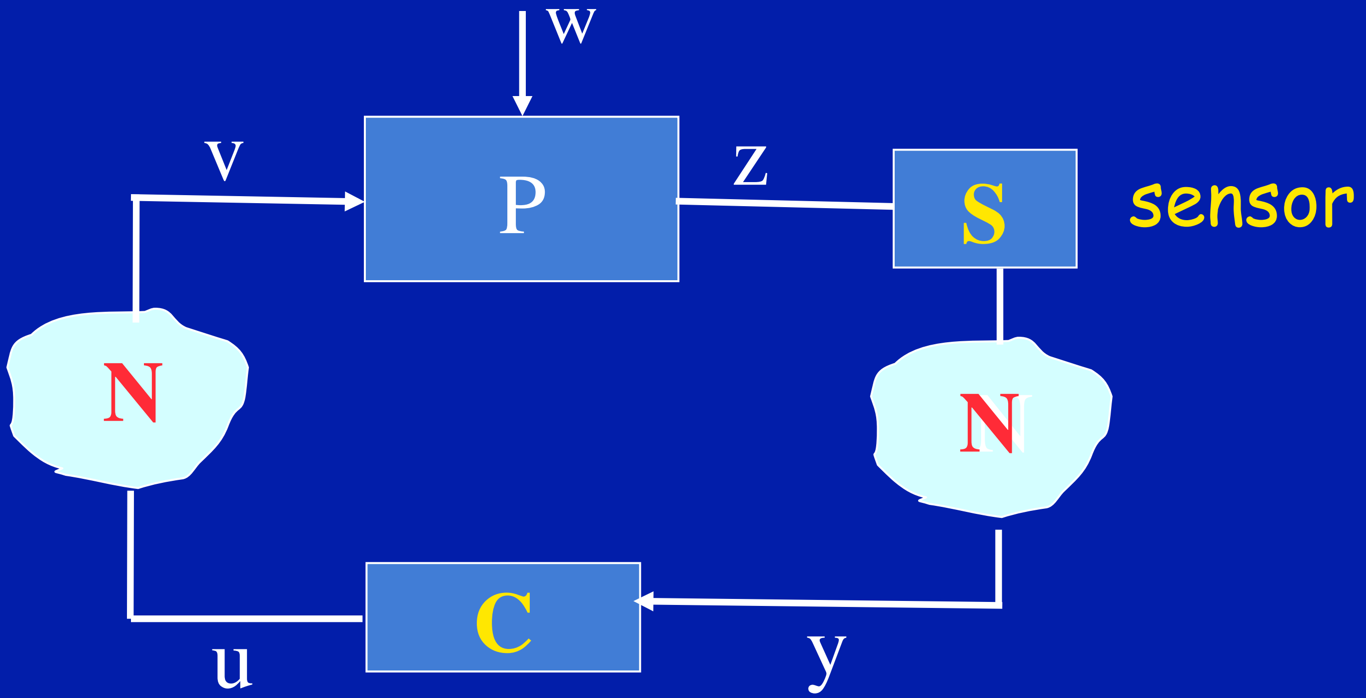
Yantai, China

July 22-24, 2011

Outline

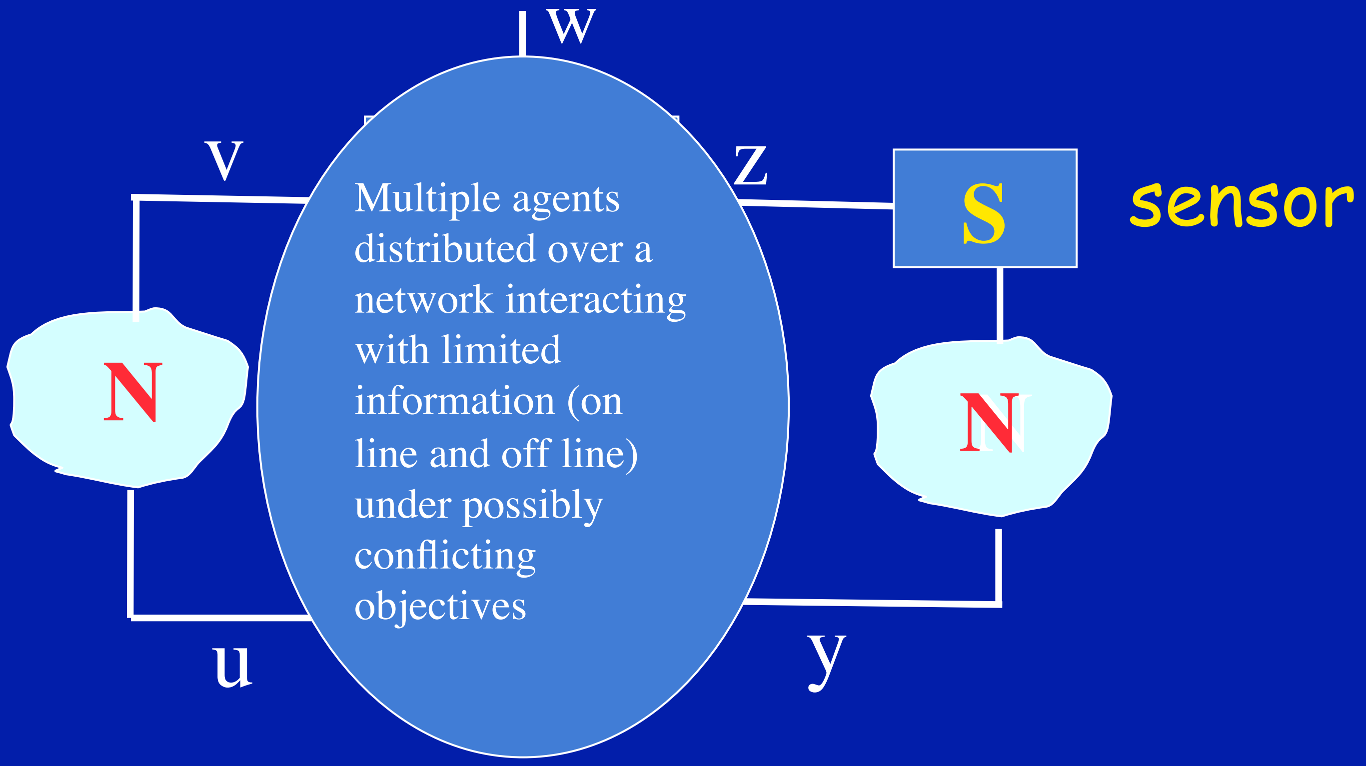
- Networks, Sensing, Control, and Jamming
- What/when/how to sense, transmit, and control with limited opportunities
- Non-classical information in multi-agent DM
- Coping with unreliability partially caused by adversarial action
- Worst-case disruption strategies and corresponding control policies
- Conclusions

Variations around the Common Paradigm
Networked Control System



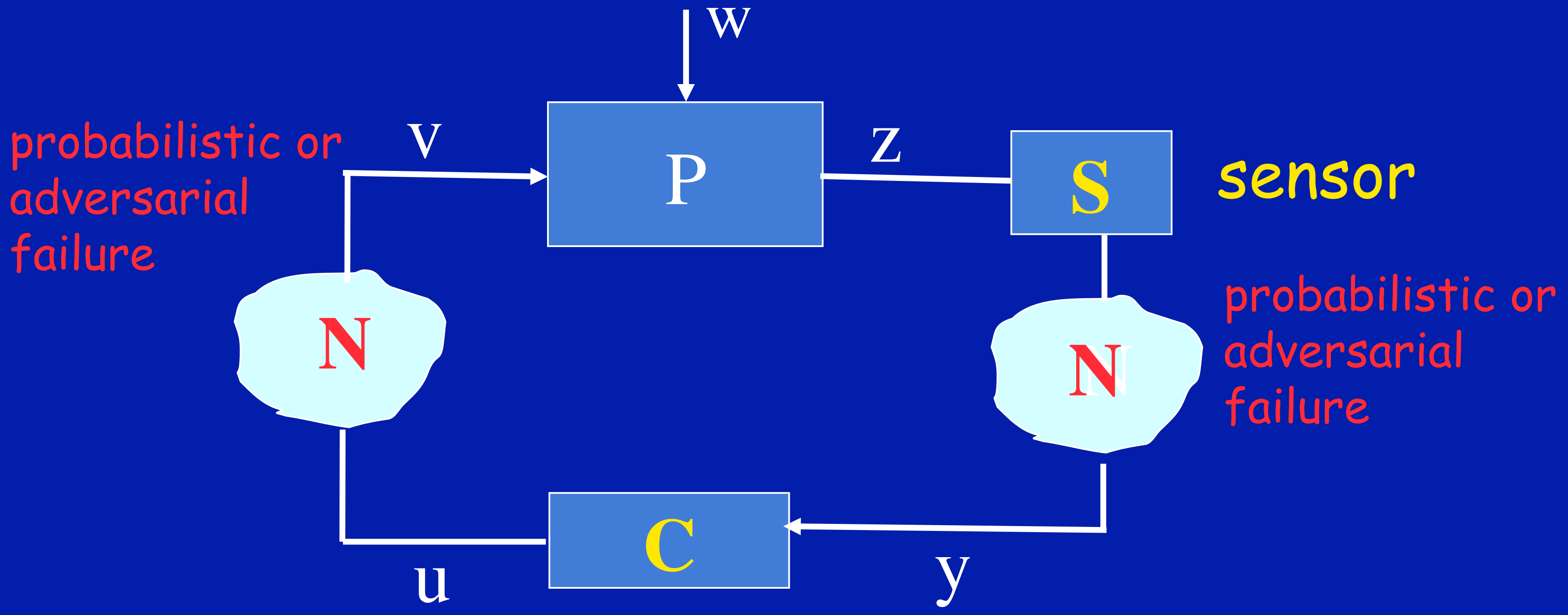
PI → optimize

Variations around the Common Paradigm
Networked Control System



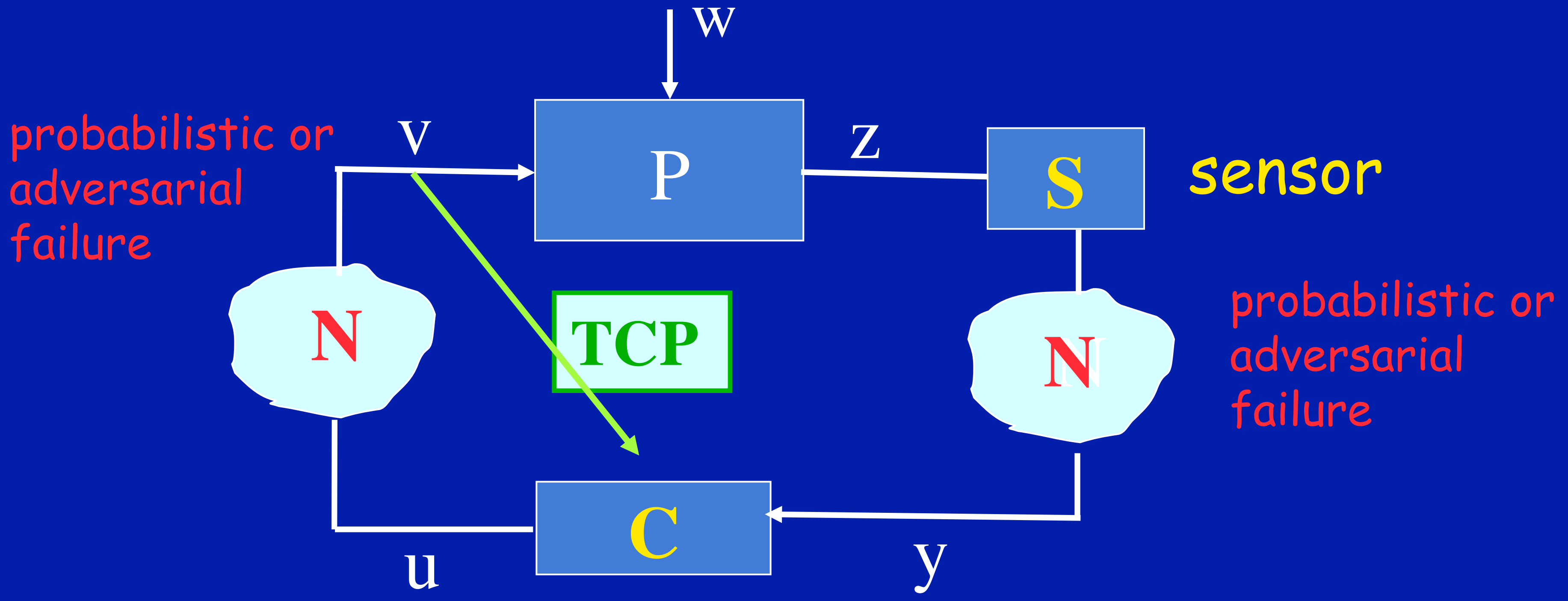
PI $\rightarrow$ optimize

Failure of Channels



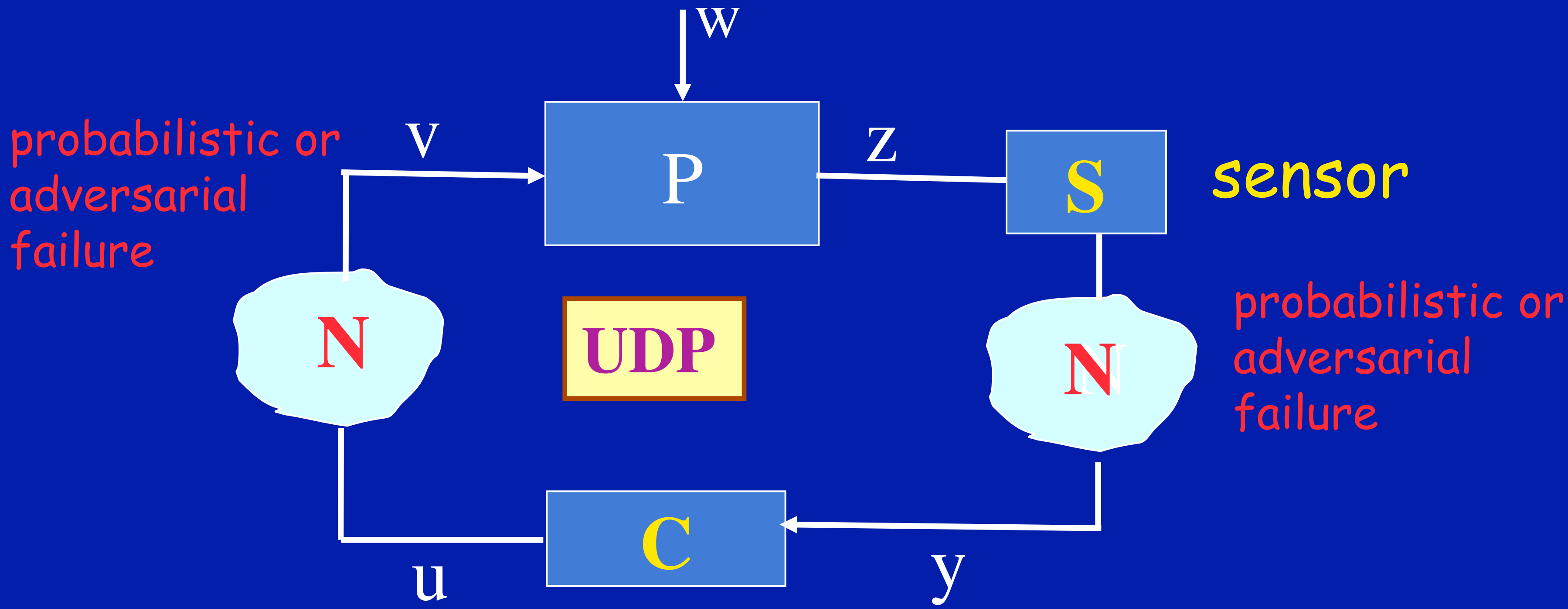
PI $\rightarrow$ optimize

Acknowledgement: Scenario I



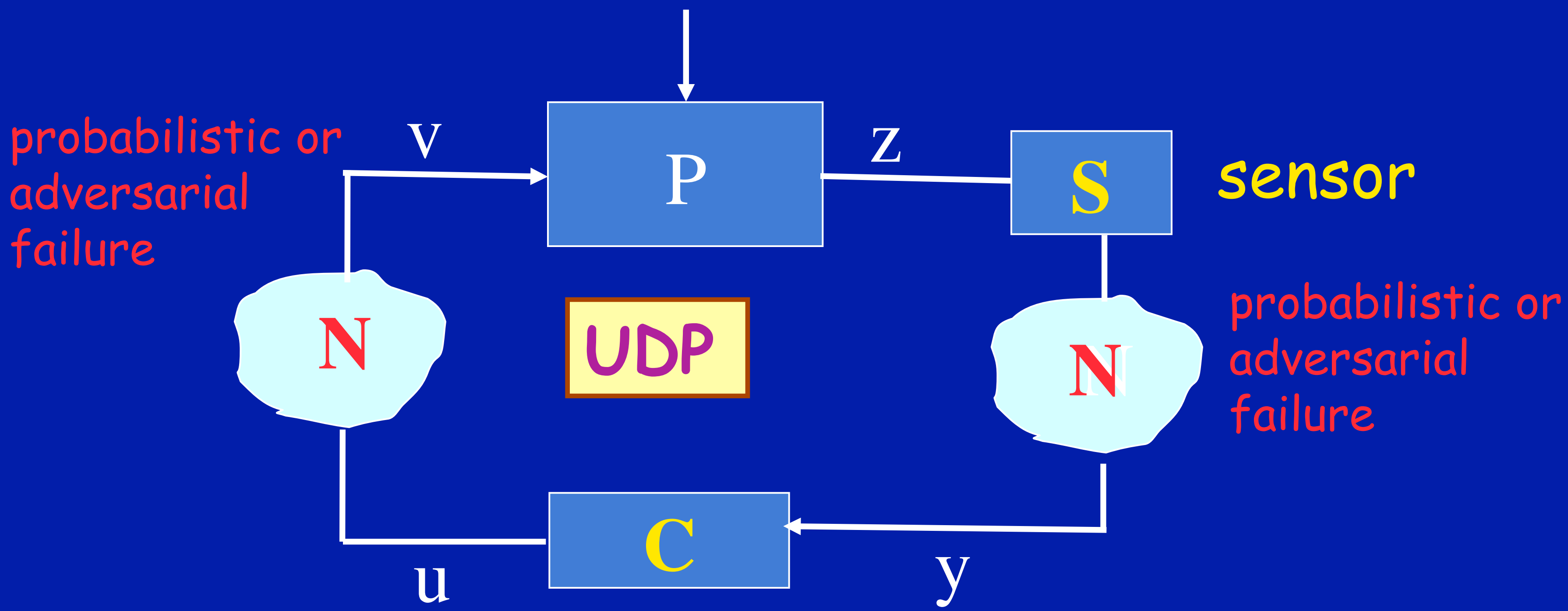
Transmission Control Protocol of Internet
(with acknowledgements)

No acknowledgement: Scenario II



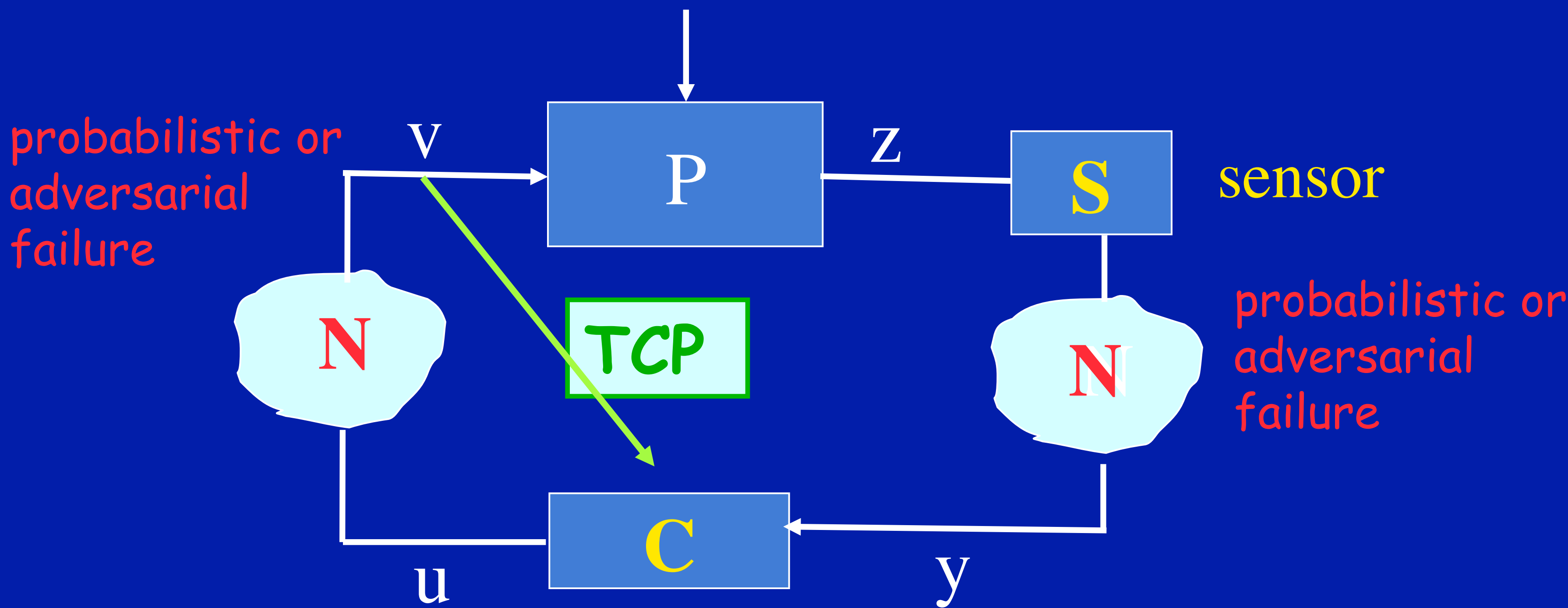
<u>U</u>ser <u>D</u>atagram <u>P</u>rotocol -- best effort network
(no acknowledgements)

Scenario II: Information structure



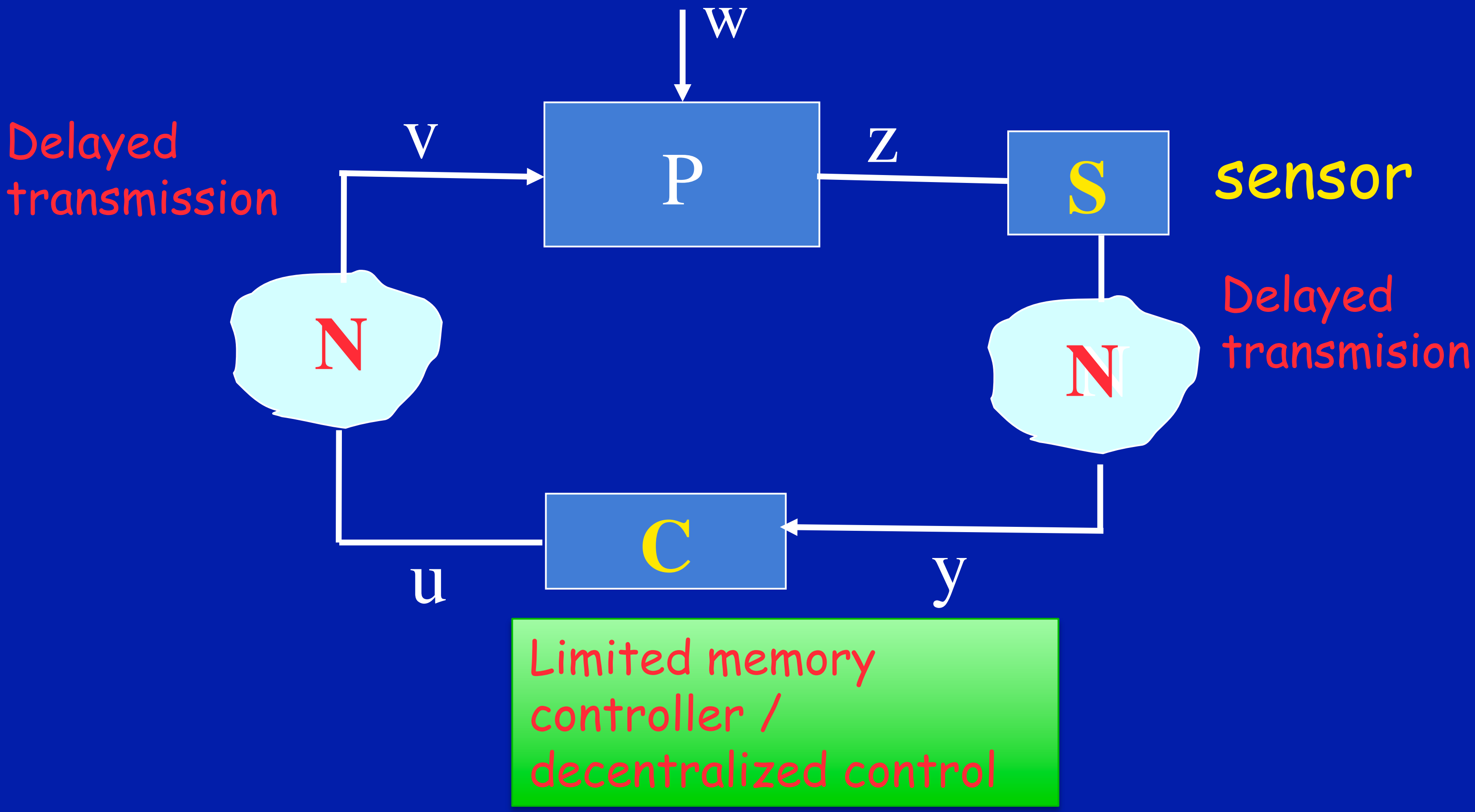
$$I_k^{UDP} = \{y_0^k, u_0^{k-1}, \beta_0^k\}$$
$$u_k = \mu_k (I_k^{UDP})$$

Scenario I: Information structure

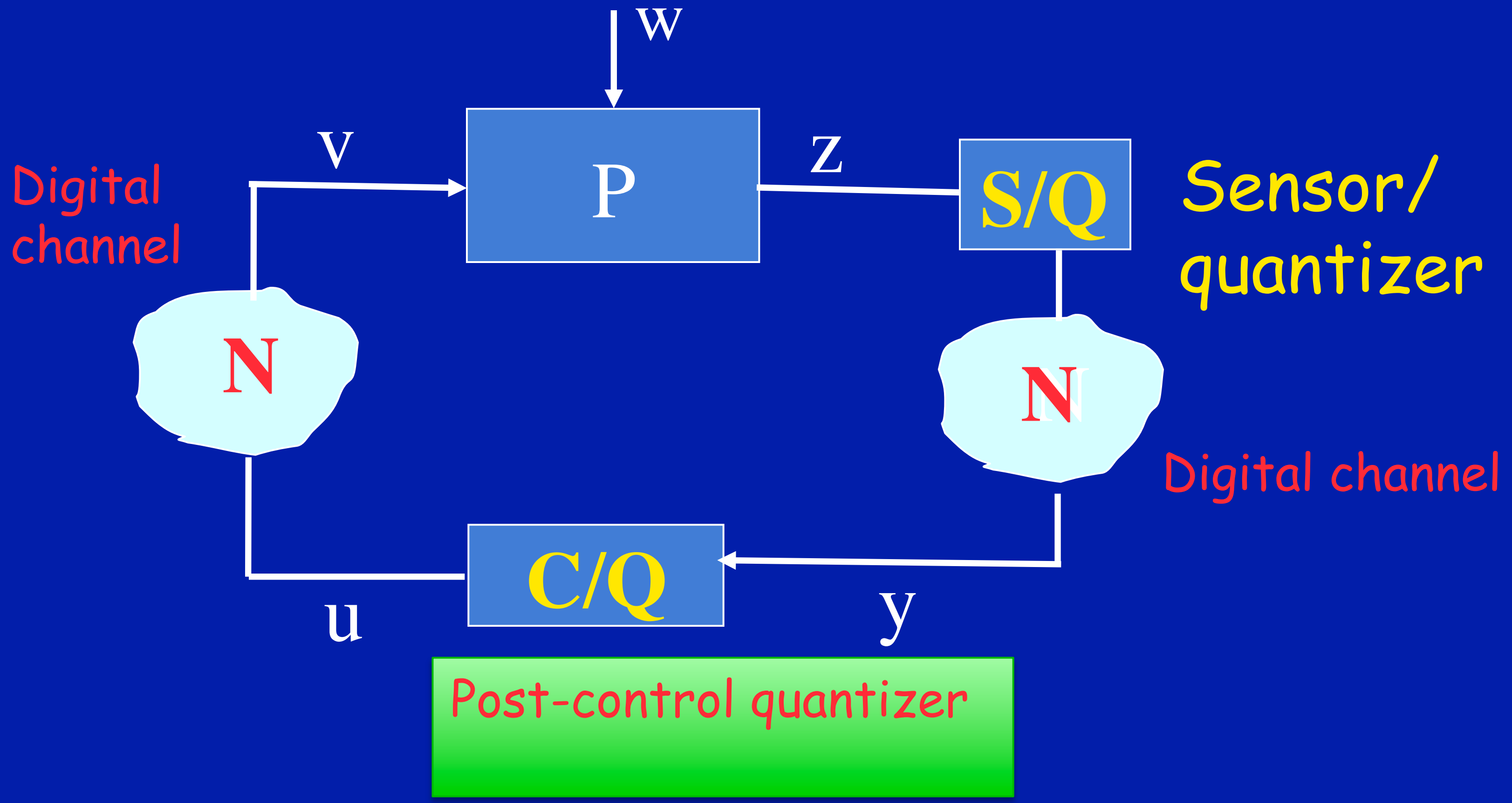


$$I_k^{TCP} = \{I_k^{UDP},\ \alpha_0^{k-1}\}$$
$$u_k = \mu_k\,(I_k^{TCP})$$

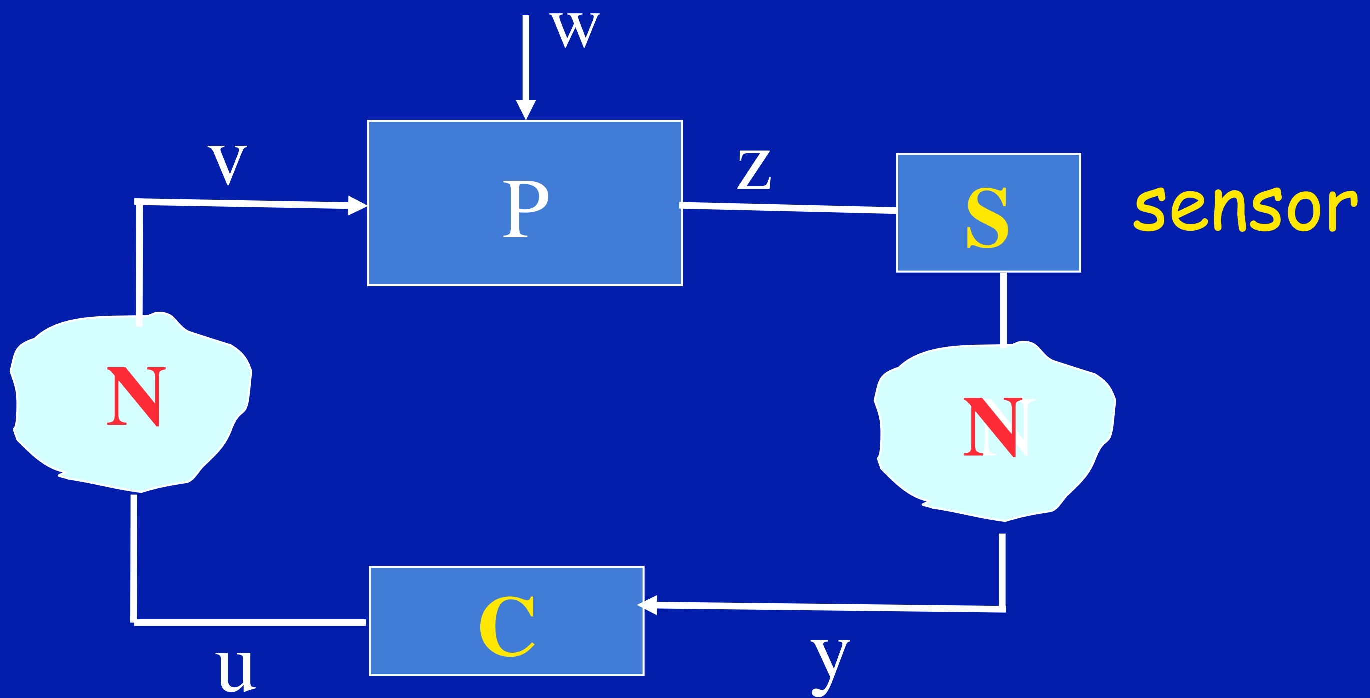
Delay, Limited Memory, Decentralization



Digital Channels: Quantization



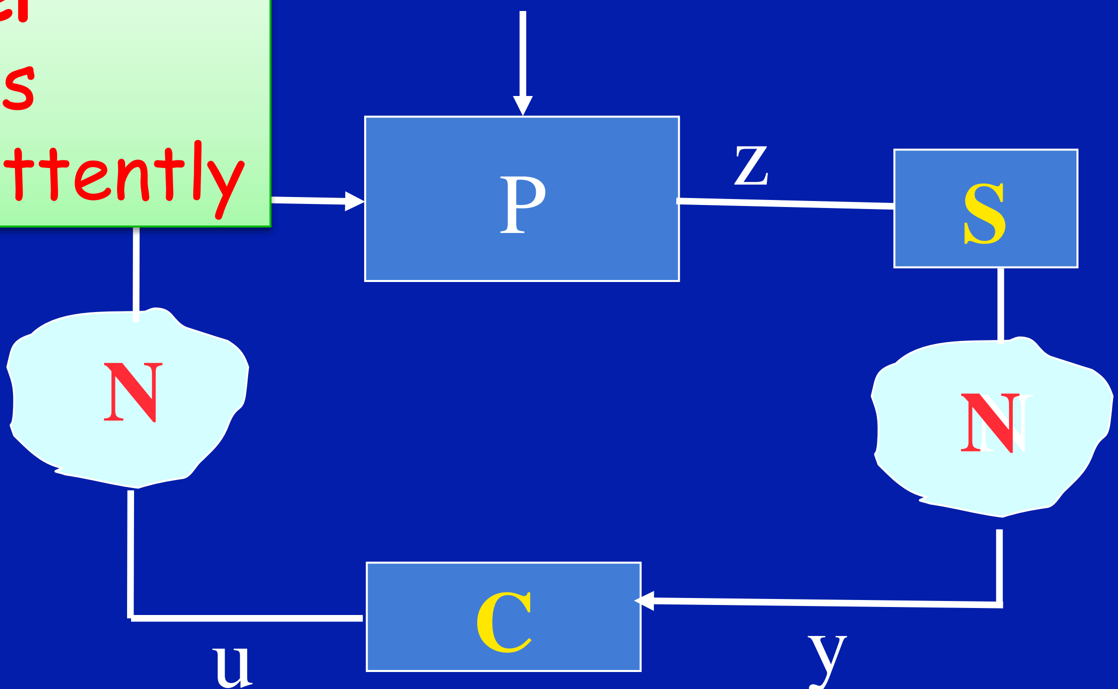
Multiple Criteria / Games



Non-cooperating controllers /multiple objectives

Limited Usage/Action

Jammer disrupts intermittently



sensor

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

Controller communicates with **Plant** intermittently

PI → optimize

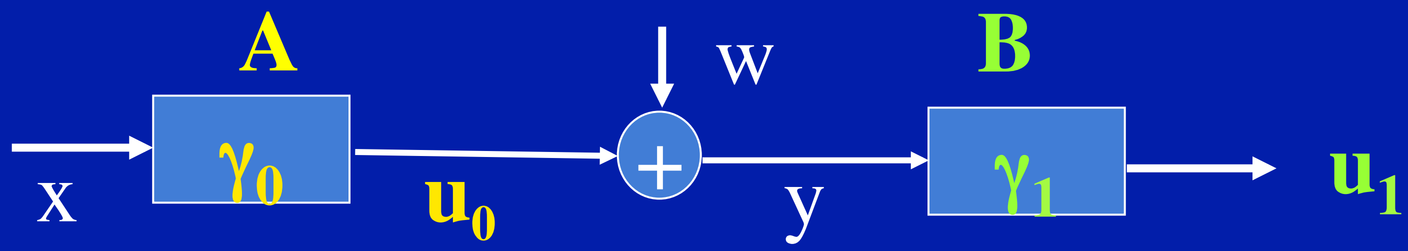
General Framework

- Multiple decision makers (agents, players) picking policies (decision laws, strategies) leading to actions that evolve over time
- Policies are constructed based on information received (active as well as passive) and guided by individual utility or cost functions over the DM horizon
  - Single DM => stochastic control
  - Single objective => stochastic teams
  - Otherwise ZS or NZS games, with NE

Coupling of Information and Actions

- Is the quality of active and relevant information received by an agent affected by actions of other agents ?
  - If no => the problem is generally "simple"
  - If yes => it is generally "difficult"

Non-classical (limited memory)

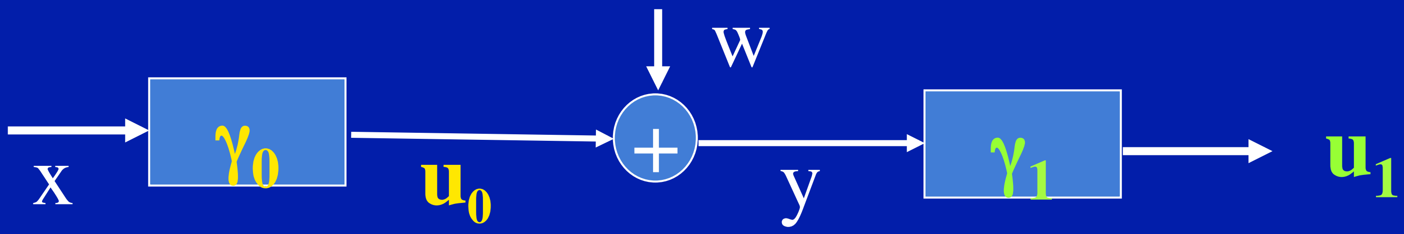


$x \sim N(0, \sigma_x^2)$ $w \sim N(0, \sigma_w^2)$

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

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

The "simplest" difficult problem
(Witsenhausen, 1968)

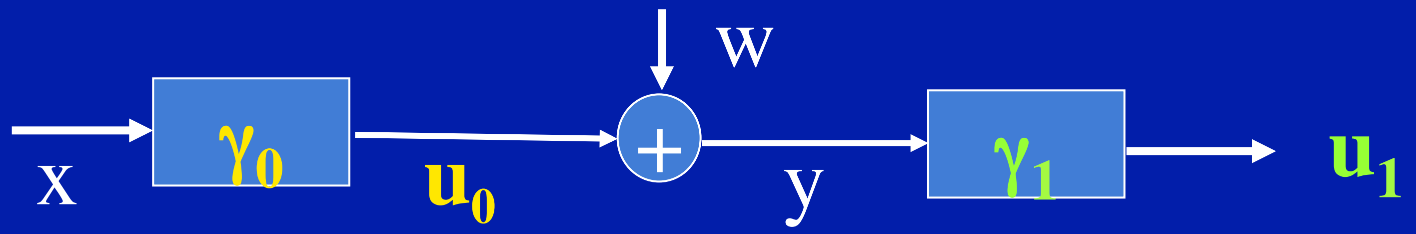


$$Q_W(x, u_0, u_1) = k_0\,(u_0 - x)^2 + (u_0 - u_1)^2$$

⇒ optimal team solution exists, but its structure is not known

affine policies are **not** optimal

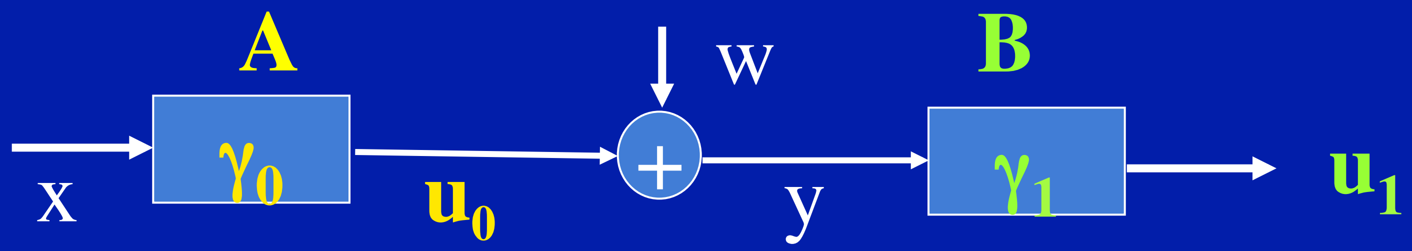
The "simplest" difficult problem
(Witsenhausen, 1968)



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

When restricted to affine policies, there exist multiple local optima.
Convexity is lost when restrictions are placed on memory.

The "simplest" difficult problem
(Witsenhausen, 1968)

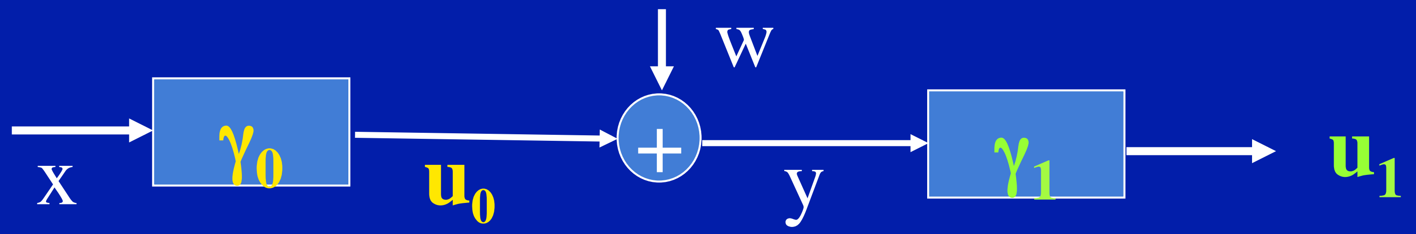


$x \sim N(0, \sigma_x^2)$ $\quad$ $w \sim N(0, \sigma_w^2)$

Note: This is a standard 2-stage DT LQG control problem except that control at stage 2 does not have access to what control at stage 1 had (*memoryless* controllers)

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

The "simplest" difficult problem



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

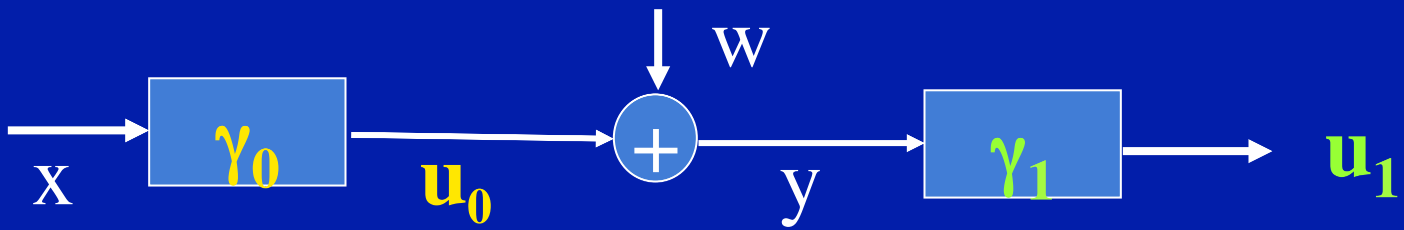
A policy pair that beats the best linear one:

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

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

optimize wrt ε and λ

The "simplest" difficult problem



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

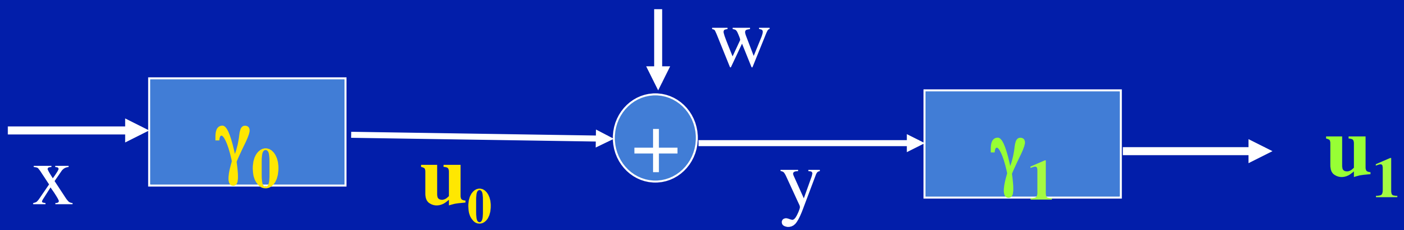
A more general one with improved perf:

$u_0 = \gamma_0(x) = \varepsilon \, \text{Quant}(x) + \lambda x$

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

optimize wrt ε and λ for different Quant(ization) schemes (Bansal, TB)

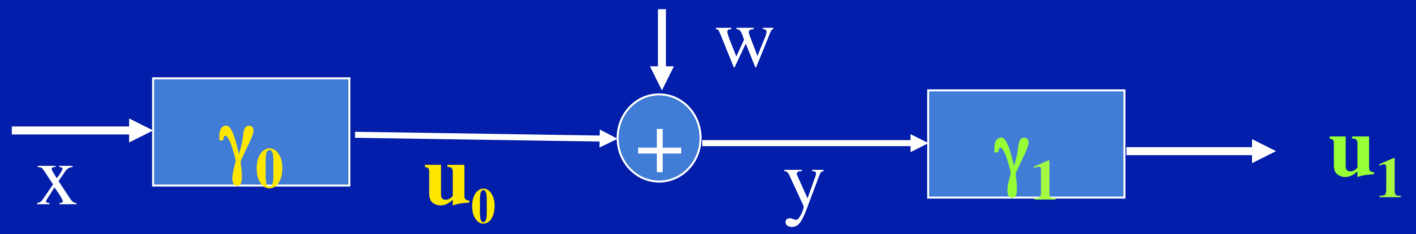
Gaussian Test Channel



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

⇨ optimal pair of decision laws (encoder/decoder) exists, and they are linear

However, with Conflicting Objectives

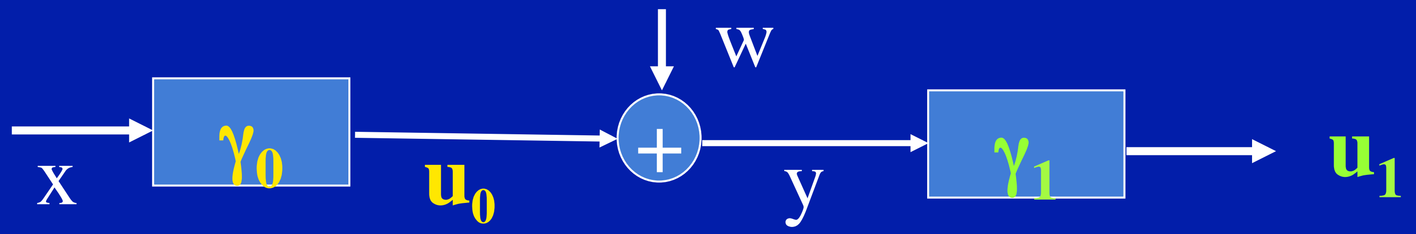


$$Q_G(x, u_0, u_1) = - k_0 (u_0 - x)^2 + (u_0 - u_1)^2$$

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

Unique saddle-point solution, policies are linear (TB)

Recap



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

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

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

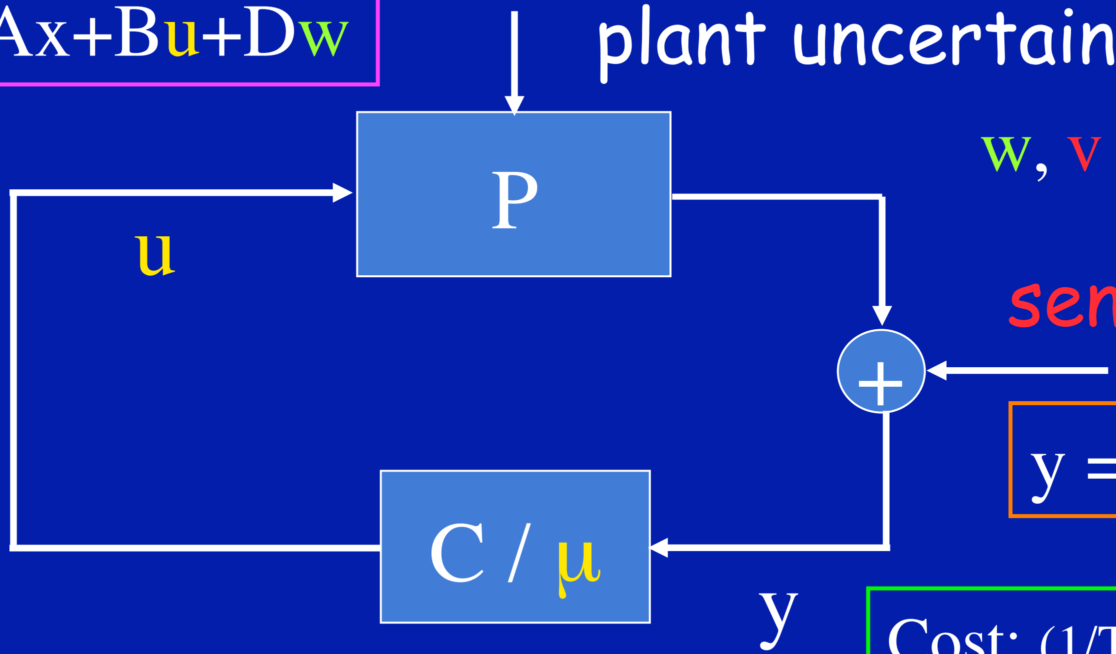
→ Not only the IS, but also the cost function is a determining factor

A message to take on a variation to LQG

$dx/dt = Ax+Bu+Dw$

plant uncertainty, w

w, v indep GWN



sensor noise v

$y = H\,x + v$

Cost: $(1/T) \int |z|^2 + |u|_R{}^2$

$u(t)$ $= \mu(y_t)$ -- memoryless controller

a difficult problem; solution not known

Another Message

Quantization plays an important role in the construction of policies that improve upon best linear ones (even though the channels are not discrete)

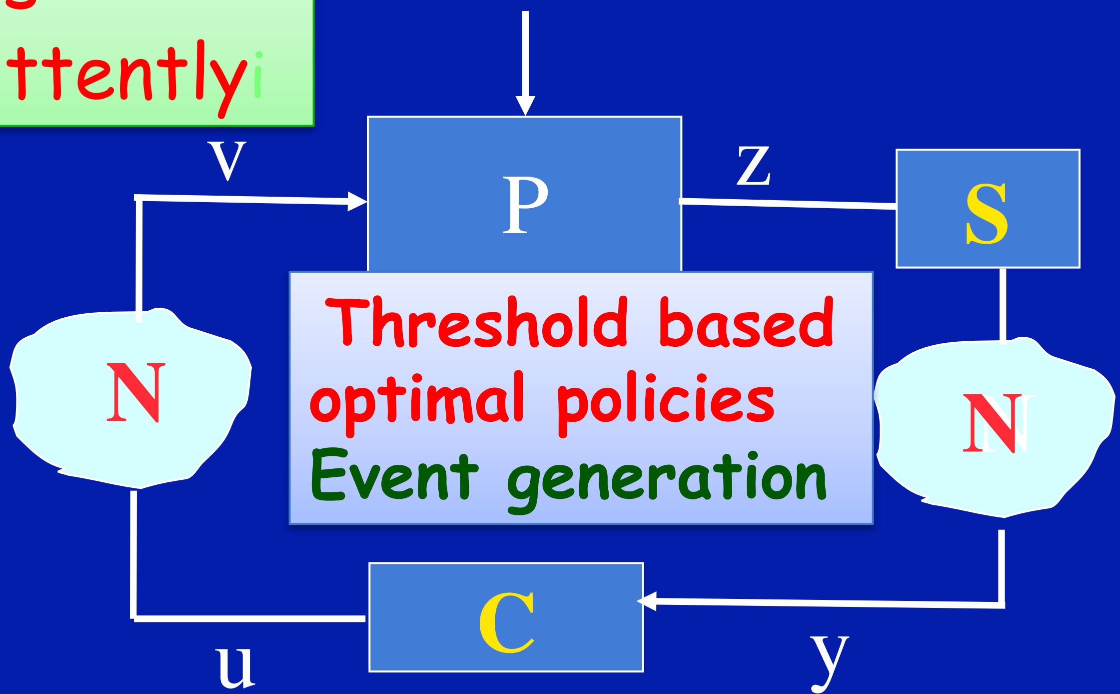
(Bansal-TB, TB, Yuksel-Tatikonda, Grover-Sahai, Lipsa-Martins)

Limited Actions and Jamming

what/when/how to transmit, control, and jam with **limited opportunities**

Limited Usage

Jammer disrupts intermittently



sensor

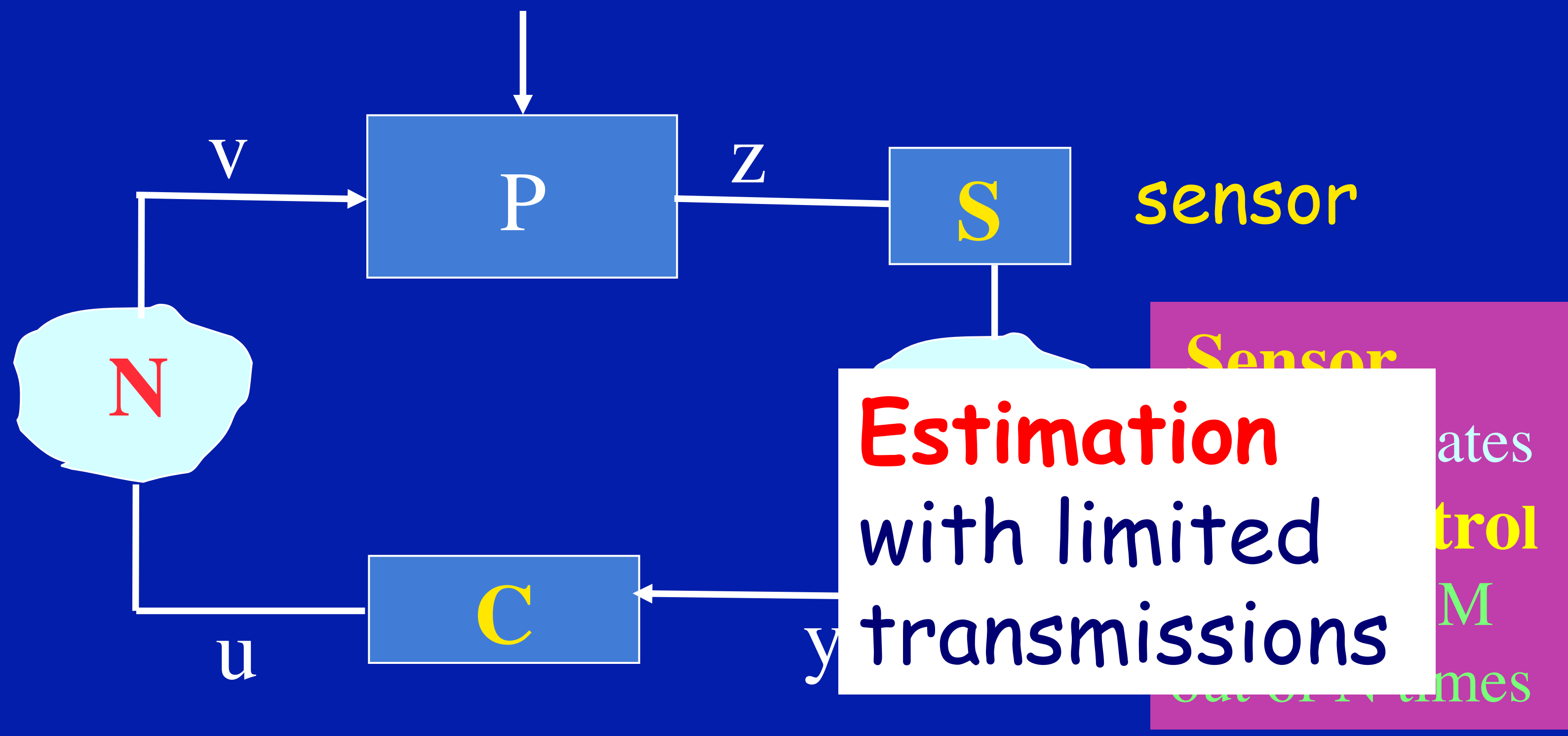
Threshold based optimal policies
Event generation

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

Controller communicates with **Plant** intermittently

PI → optimize

Limited Usage

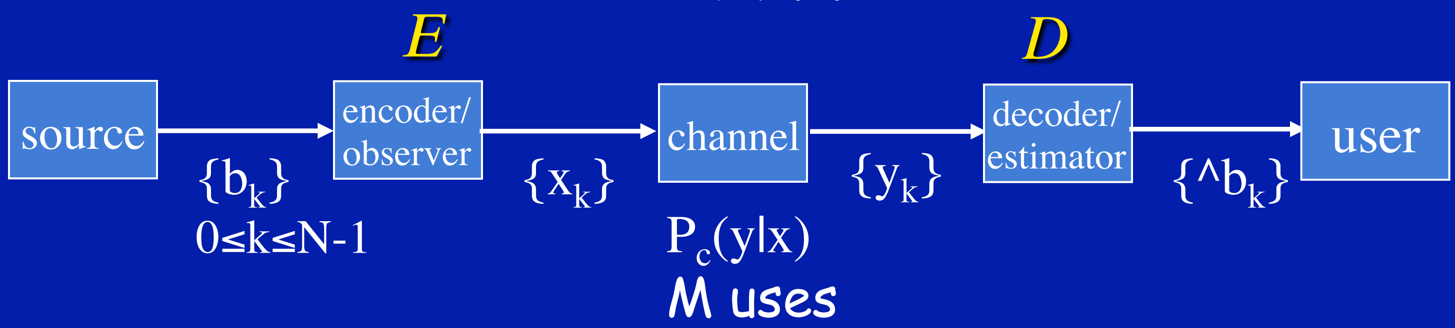


Estimation with limited transmissions

PI → optimize

(Imer-TB'10)

Optimal Estimation over a Limited-Use Channel

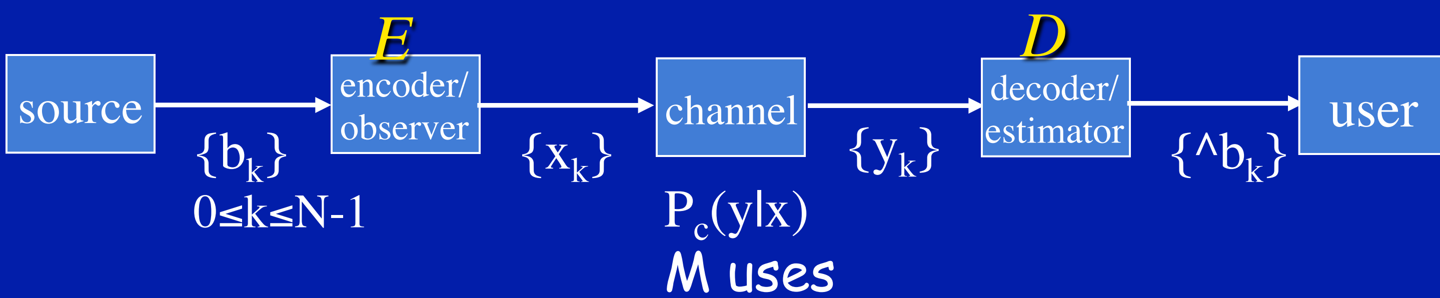


$x_k = \mathcal{E}(z_k)$ $\quad x \in \mathcal{X}$ $\quad y \in \mathcal{Y}$

$z_k = b_k + v_k$ $\quad M < N$

Given a "source" and a "memoryless channel", for a given message length N, and number of channel uses M, and with some power constraint on the encoder, what is the minimum attainable value of the average distortion $D_{(M,N)}$ and a corresponding E & D pair?

Optimal Estimation over a Limited-Use Channel



$x_k = E(z_k)$

$z_k = b_k + v_k$

$x \in X \quad y \in Y$

$M < N$

Dynamic & Non-classical

Order of actions at time k:

1. b_k (or z_k) becomes available to the sensor
2. Sensor makes a decision: transmit/shape or not
3. Estimator acts by generating $\hat{b}_k$
4. Estimation error is incurred and we move to k+1

A Special Case: 2 string, *iid*, no noise

$N=2$, $M=1$, b_0, b_1 i.i.d. Gaussian, 0-mean, variance σ^2

Perfect channel, no noise

Estimation error: $e = E\{(b_0 - \hat{b}_0)^2 + (b_1 - \hat{b}_1)^2\}$

Open-loop sensor policy:

Arbitrarily picks transmission time ==> $e_{OL} = \sigma^2$

Closed-loop sensor policy:

Transmit b_0 if it lies outside $[\alpha, \beta]$, $\alpha < 0 < \beta$; otherwise b_1

Minimization problem faced by sensor:

$$e_{(\alpha,\beta)} = \int_\alpha^\beta (b - E[b \mid b \in [\alpha, \beta]])^2 f(b)\, db + \sigma^2 P\{b_0 \notin [\alpha, \beta]\}$$

Special Case: Solution

$$(\alpha^*, \beta^*) = (-\sigma, \sigma)$$

$\Rightarrow$

$$e_{CL}{}^* = e_{(\alpha^*, \beta^*)} = [1 - \sqrt{(2 / \pi e)}]\, \sigma^2$$
$$\approx 0.52\, \sigma^2$$

$\Rightarrow$

48% improvement over the OL policy

Special Case: Solution

$(\alpha^*, \beta^*) = (-\sigma, \sigma)$

$\Rightarrow$

The knowledge of no action is useful information !!

$\Rightarrow$

48% improvement over the OL policy

General Solution for linear systems (Imer, TB)

Best sensor policy is of **threshold form**:
At time k transmit z_k if it is in a measurable set ¥(s_k,t_k), otherwise do not
¥(s,t) obtained offline as the minimizer in a recursive equation satisfied by accumulated optimum error, $e^*(s,t)$, at each point (s,t):

$$e^*(s,t) = \min_{¥(s,t)} \{ e^*(s-1, t-1)\, \mathrm{Prob}(z \in ¥) + e^*(s, t-1)\, \mathrm{Prob}(z \notin ¥) + \text{average error at } (s,t) \text{ due to decision at } (s,t) \}$$

$e^*(t,t) = 0$, $e^*(0,t) = t$ var(of input rv)

Specific structure of ¥(s,t) depends on the pdf/pmf and PI.

Explicit Solution in a Special Case

Continuous distribution f for b, f(-b) = f(b)

No noise v (from source to sensor) -- n.l.o.g

==> $¥^c(s,t) = [-\beta_{(s,t)}, \beta_{(s,t)}]$

$$\beta_{(s,t)} = \sqrt{\{e^*(s-1, t-1) - e^*(s, t-1)\}}$$

Gaussian: $\varepsilon_{(s,t)} := e^*(s,t) / \mathrm{var}(b)$

$$\varepsilon_{(s,\, t)} = \varepsilon_{(s-1,\, t-1)} - [(\beta_{(s,t)})^2 - 1][2\Phi(\beta_{(s,t)}) - 1] - (2/\sqrt{2\pi})\, \beta_{(s,t)} \exp\left((\beta_{(s,t)})^2 / 2\right)$$

$$\varepsilon_{(t,\, t)} = 0, \quad \varepsilon_{(0,\, t)} = t$$

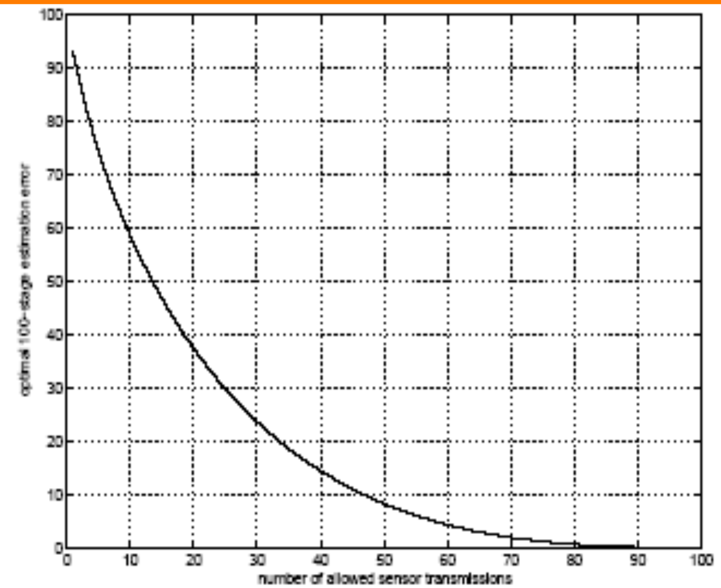
An Illustrative Example

Problem: Given a time-horizon of length N=100, estimate the state of a zero-mean *i.i.d.* Gaussian process with unit variance

Design Criterion: The cumulative estimation error should not exceed 20.

Solution I: Make 80 sensor transmissions picked at arbitrary times

Solution II: Use the optimal sensor transmission and estimation policies

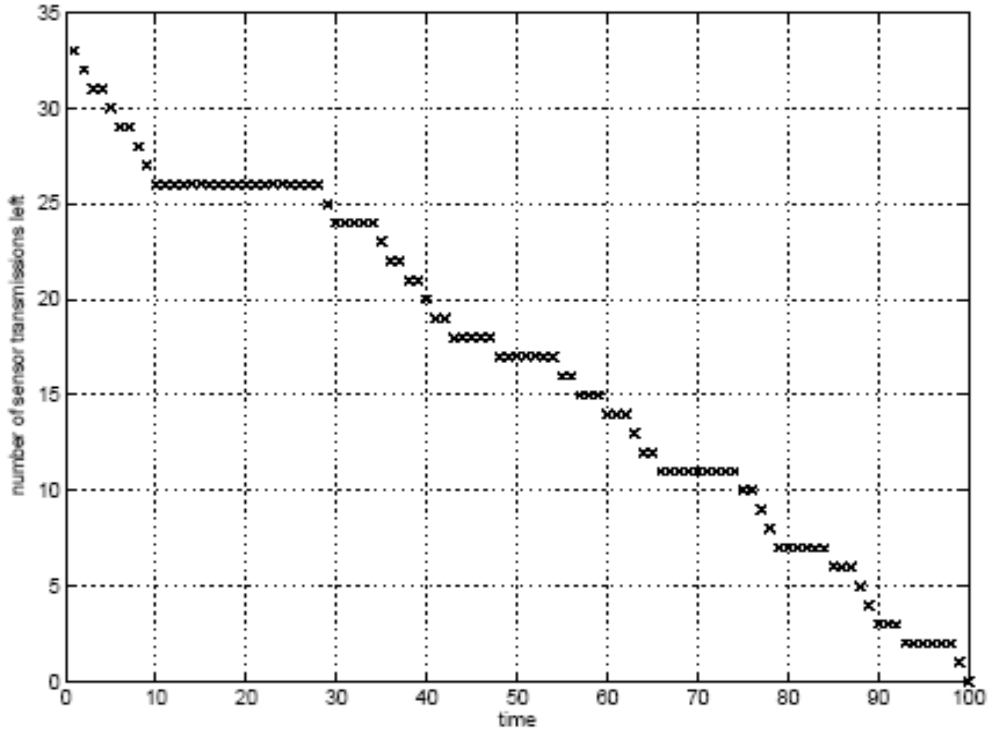


An Illustrative Example (cont.)

Estimation error of 20 can be achieved with 34 transmissions!

This is approximately 58% improvement --- considerable savings in battery power (if sensor is power-limited) or transmission slots (if the sensor is time-slot limited.

An Illustrative Example (cont.)



Typical sample path of the number of sensor transmissions left under the optimal transmission policy of the sensor

$$(N, M) = (100, 34)$$

Source as a Markov Process

$b_{k+1} = A\, b_k + w_k$ $\{w_k\}$ GWN

Optimum sensor policy: keeps track of 3 variables (r_k, s_k, t_k)

r_k : # time units passed since last transmission

At time k transmit b_k if it is in a measurable set ¥(r_k, s_k, t_k), otherwise do not.

¥(r,s,t) obtained offline as the minimizer in a recursive equation satisfied by accumulated optimum error, $e^*(r,s,t)$, at each (r,s,t):

$$e^*(r,s,t) = \min_{¥(r,s,t)} \{ e^*(1, s-1, t-1)\, \mathrm{Prob}(b_{N-t} \in ¥) + e^*(r+1, s, t-1)\, \mathrm{Prob}(b_{N-t} \notin ¥) + \text{average error at } (r,s,t) \text{ due to decision at } (r,s,t) \}$$

Again an interval solution

$¥^c(r,s,t) = [\alpha_{(r,s,t)}, \beta_{(r,s,t)}]$

$\beta_{(r,s,t)} = A^r b_{N-t-r} + \sqrt{\{e^*(1,s-1, t-1) - e^*(r+1,s, t-1)\}}$

$\alpha_{(r,s,t)} = A^r b_{N-t-r} - \sqrt{\{e^*(1,s-1, t-1) - e^*(r+1,s, t-1)\}}$

$\varepsilon_{(r,s,t)} := e^*(r,s,t) / \sum_{k=1}^{r} A^{2(k-1)} var(b_0)$

$\varepsilon_{(r,s, t)} = \varepsilon_{(1,s-1, t-1)} - [(\nu_{(r,s,t)})^2 -1][2\Phi(\nu_{(r,s,t)}) -1]$

$-(2/\sqrt{2\pi})\ \nu_{(r,s,t)} \exp ((\nu_{(r,s,t)})^2 / 2)$

$\nu_{(r,s,t)} := \sqrt{\{e^*(1,s-1, t-1) - e^*(r+1,s, t-1)\}}$

Multi-StepMarkov Process

$b_{k+1} + a_0\, b_k + \ldots + a_{n-1}\, b_{k-n+1} = w_k \qquad \{w_k\}$ GWN

Optimum sensing again keeps track of 3 variables (r_k, s_k, t_k)

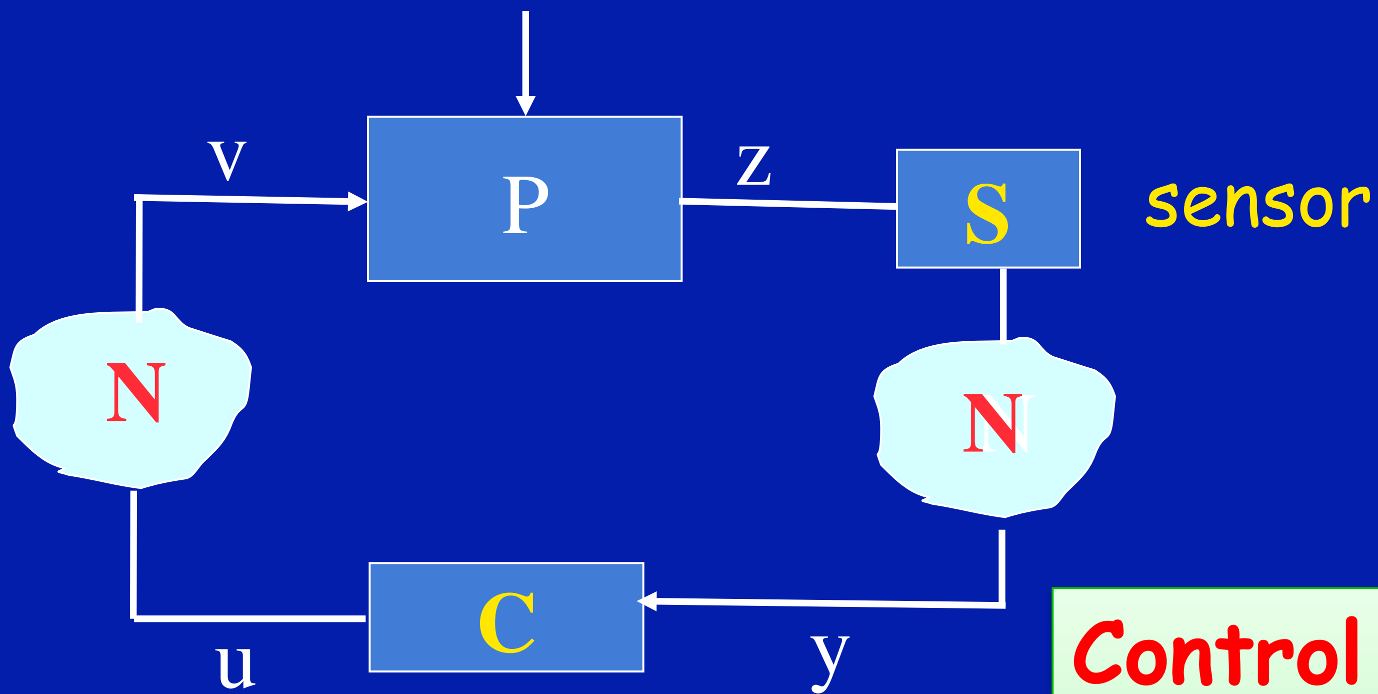
r_k : # time units passed since last transmission

At time k transmit b_k if it is in a measurable set ¥(r_k, s_k, t_k), otherwise do not.

¥(r,s,t) obtained offline as the minimizer in a recursive equation satisfied by accumulated optimum error, $e^*(r,s,t)$, at each (r,s,t):

$$e^*(r,s,t) = \min_{¥(r,s,t)} \{ e^*(1, s-1, t-1)\, \mathrm{Prob}(b_{N-t} \in ¥) + e^*(r+1, s, t-1)\, \mathrm{Prob}(b_{N-t} \notin ¥) + \text{average error at } (r,s,t) \text{ due to decision at } (r,s,t) \}$$

Limited Usage



PI → optimize

Controller communicates with **Plant** intermittently

Control with limited transmissions M out of N

General Solution for linear-quadratic systems
(Imer, TB)

Best **control** policy is of **threshold form**:

At time k generate a control signal u_k and transmit it if the state or its conditional mean is in a measurable set ¥$_C(s_k,t_k)$; otherwise do not

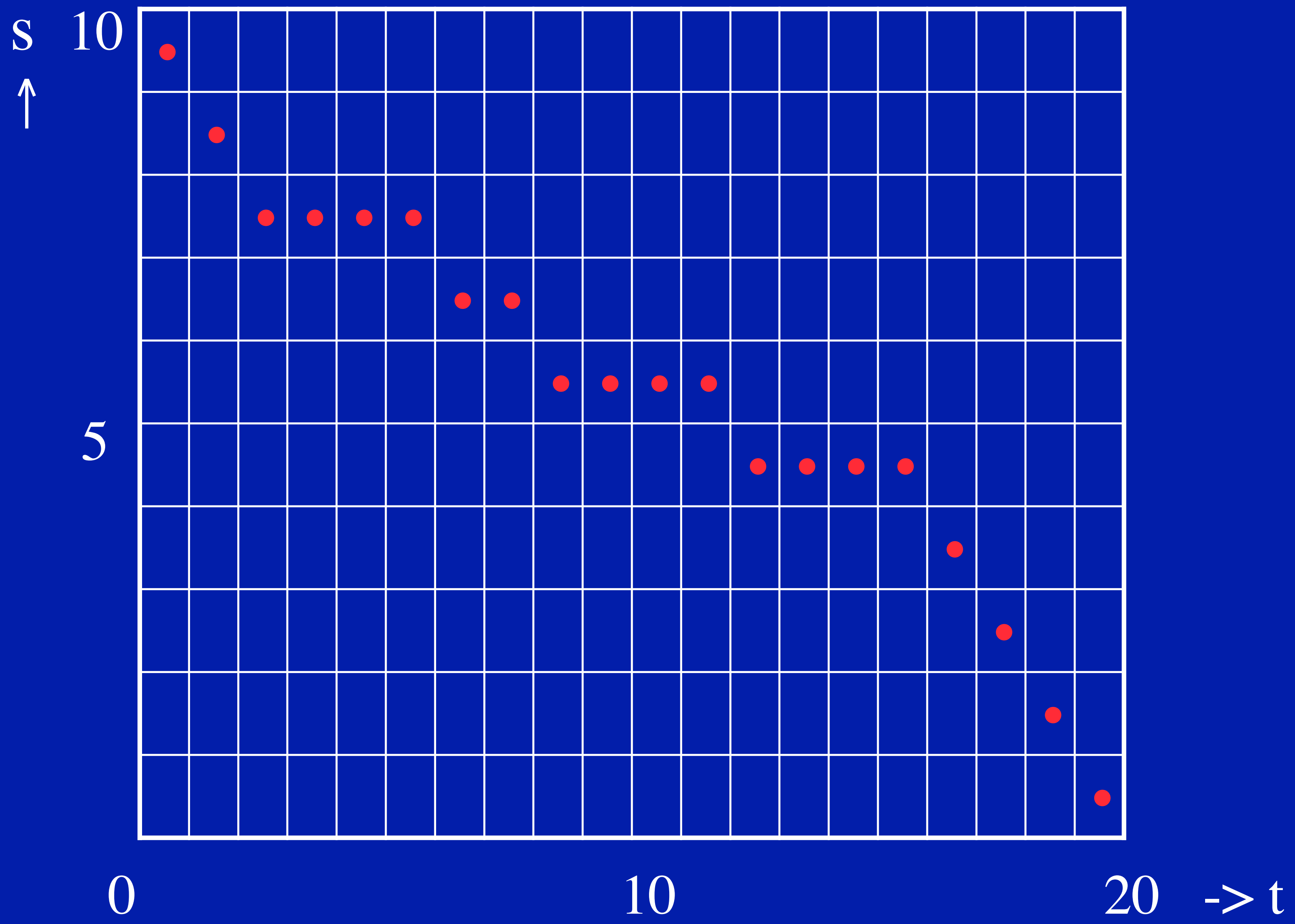
Specific structure of ¥$_C(s_k,t_k)$ depends on the pdf/pmf of system and channel noises, whether control-plant communication is noisy, and also on PI.

Numerical Solutions

Numerical integration was used to compute the recursions $\Delta_{(s,t)}$, which led to thresholds $\tau(s, t)$

Implemented the optimal control with M actions for an N-stage problem (N=20). Computed $J_{(M,N)}^*$ based on sample paths, and for different M values

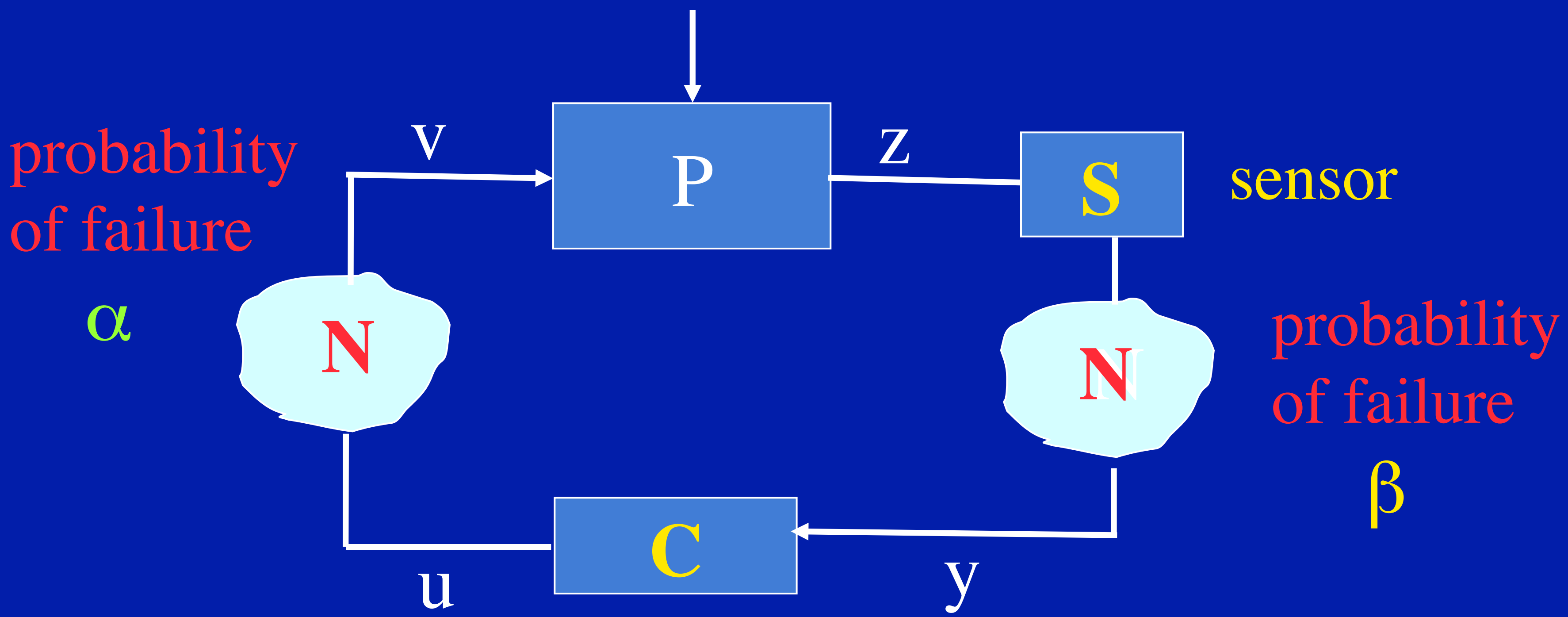
Times of control action ==>


s
10
5
0
10
20
-> t

M	$J^*_{(M,N)}$	%
1	96.4266	203.9327
2	68.1907	114.9343
3	47.2060	48.7914
4	44.0160	38.7366
5	39.8642	25.6503
6	37.1557	17.1132
7	35.6168	12.2627
8	34.1551	7.6555
9	33.6935	6.2005
10	33.6913	6.1936

M	$J^*_{(M,N)}$	%
11	33.2445	4.7853
12	32.9262	3.7820
13	32.8267	3.4684
14	32.4936	2.3249
15	32.1082	1.2037
16	31.9824	0.8072
17	31.8822	0.4914
18	31.8417	0.3637
19	31.7337	0.0233
20	31.7263	0

Link Failures / Lossy Transmission



Lossy Transmission

$x_{k+1} = f(x_k, v_k, w_k), \quad k = 0,1,\ldots$

$v_k = \alpha_k u_k$ or $v_k = v_{k-1}$ if $\alpha_k = 0$

$y_k = \beta_k z_k \qquad z_k = h(x_k, w_k)$

$\{\alpha_k\}, \{\beta_k\}$ independent *i.i.d.* Bernoulli

$\text{Prob}(\alpha_k = 0) = \alpha, \qquad \text{Prob}(\beta_k = 0) = \beta$

$\{w_k\}$ *i.i.d.* plant / channel noise

Lossy Transmission

$x_{k+1} = f(x_k, v_k, w_k), \quad k = 0,1,\ldots$

$v_k = \alpha_k u_k$ or $v_k = v_{k-1}$ if $\alpha_k = 0$

$y_k = \beta_k z_k \qquad z_k = h(x_k, w_k)$

$\{\alpha_k\}, \{\beta_k\}$ independent *i.i.d.* Bernoulli

$\text{Prob}(\alpha_k = 0) = \alpha, \qquad \text{Prob}(\beta_k = 0) = \beta$

Control : $u_k = \mu_k (I_k^{TCP})$
and can act
M out of N times

Lossy Transmission

$x_{k+1} = f(x_k, v_k, w_k),\quad k = 0,1,\ldots$

$v_k = \alpha_k u_k$ or $v_k = v_{k-1}$ if $\alpha_k = 0$

$y_k = \beta_k z_k \qquad z_k = h(x_k, w_k)$

$\{\alpha_k\}, \{\beta_k\}$ independent *i.i.d.* Bernoulli

$\mathrm{Prob}(\alpha_k = 0) = \alpha, \qquad \mathrm{Prob}(\beta_k = 0) = \beta$

PI : $E_\mu \{q(x_N) + \sum_k g(x_k, v_k)\} =: J(\mu_0^N, N)$

LQG with erasure channels and limited transmissions

$$x_{k+1} = Ax_k + \alpha_k Bu_k + w_k , \quad k = 0,1,\ldots$$

$$y_k = \beta_k x_k \quad (\text{or } \; y_k = x_k + n_k) \qquad z_k = x_k$$

$$\text{Prob}(\alpha_k = 0) = \alpha, \qquad \text{Prob}(\beta_k = 0) = \beta$$

u_k is applied M times

$$J(\mu_0^N, N) = E_\mu \{|x_N|_F^2 + \textstyle\sum_k |x_k|_Q^2 + \alpha_k|u_k|_R^2\}$$

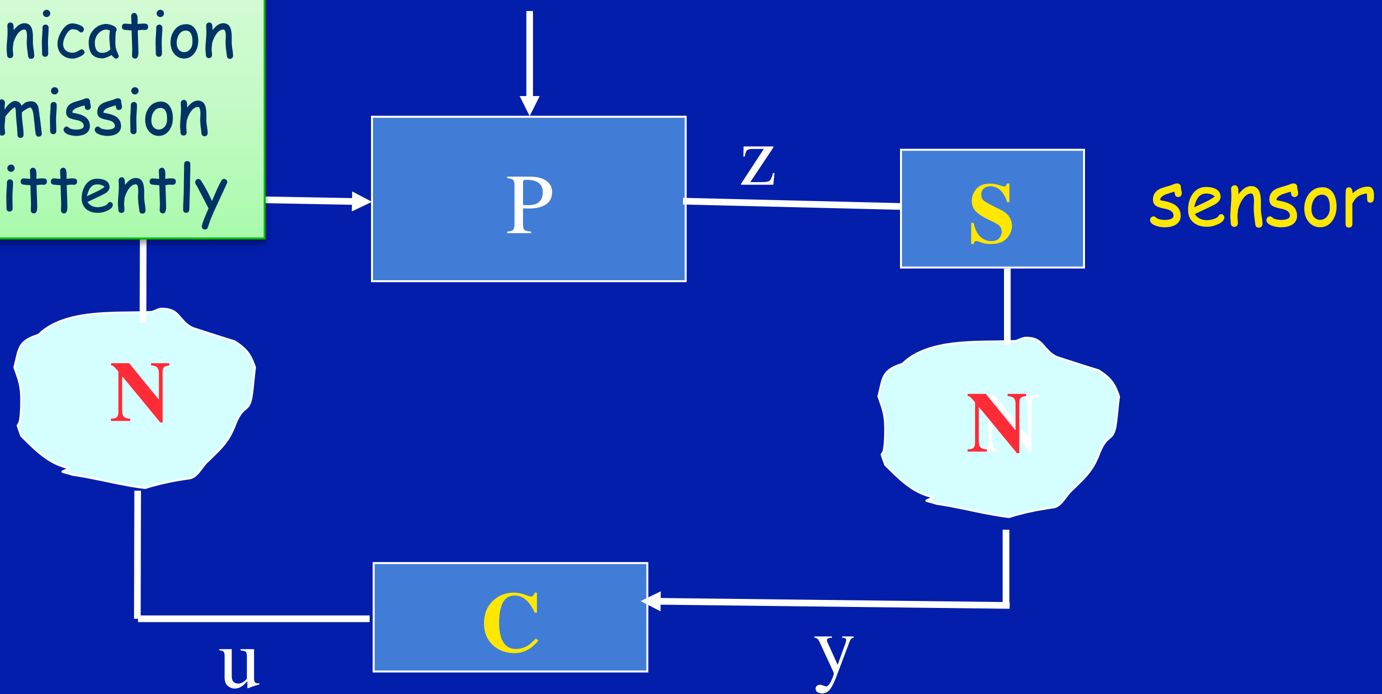
$$\text{or} \quad \lim\sup_{N\to\infty} (1/N)\, J(\mu_0^N, N)$$

Solution Summary

- Optimal cost-to-go takes into account the possibility of packet losses
- Propagation in two variables (s,t)
- Propagation of conditional mean and conditional covariance
- Decision to transmit or not based on thresholds / decision regions computed offline ($\Delta_{(s,t)} = J^0_{(s,t)} - J^1_{(s,t)} = 0$)

Limited Usage

Adversary disrupts/jams communication /transmission intermittently



PI $\rightarrow$ optimize

With Adversarial Action

(Gupta, Langbort, TB 2010)

- $x_{k+1} = Ax_k + \alpha_k u_k + w_k, \;\; k = 0, 1, \ldots, N$
- $\{\alpha_k\}$ a 0-1 variable, controlled by adversary,

  $\Sigma_{k=0}^{N-1} (1-\alpha_k) = M < N$
- $u_k = \mu_k(I_k), \quad \alpha_k = \zeta_k(I_k)$

  $I_k = \{x_{[0,k]}, \alpha_{[0,k-1]}\}, k \geq 1; \;\; I_0 = \{x_0\}$

- Cost: $E \{ \Sigma_{k=0}^{N-1} \; (x_{k+1})^2 + \alpha_k (u_k)^2 \} =: J(\mu, \zeta)$
- SP (if exists): $J(\mu^*, \zeta) \leq J(\mu^*, \zeta^*) \leq J(\mu, \zeta^*)$

With Adversarial Action

- $x_{k+1} = Ax_k + \alpha_k u_k + w_k$, $k = 0, 1, \ldots, N$
- $\{\alpha_k\}$ a 0-1 variable, controlled by adversary,

Extended state: (x, s, t)
  s = # remaining jamming instances
  $t = N-k$ (# remaining stages)
Two possible transitions from (x, s, t):
  •No jammer action: $(Ax+u+w, s, t-1)$
  •Jammer action: $(Ax+w, s-1, t-1)$

With Adversarial Action
--solution process--

- $x_{k+1} = Ax_k + \alpha_k u_k + w_k, \;\; k = 0, 1, \ldots, N$
- $\{\alpha_k\}$ a 0-1 variable, controlled by adversary,

Isaacs equation on the extended state space:

$$V_{(0,0)}(x) = x^2$$

$$V_{(s,t)}(x) = \inf_u \max_{\alpha\, \varepsilon\, \tilde{A}(s,t)} E\{x^2 + \alpha u^2 + + V_{(\check{s}(\alpha),t-1)}(Ax + \alpha u + w)\}$$

where $\check{s}(\alpha) := s$ if $\alpha = 1$

$= s-1$ if $\alpha = 0$

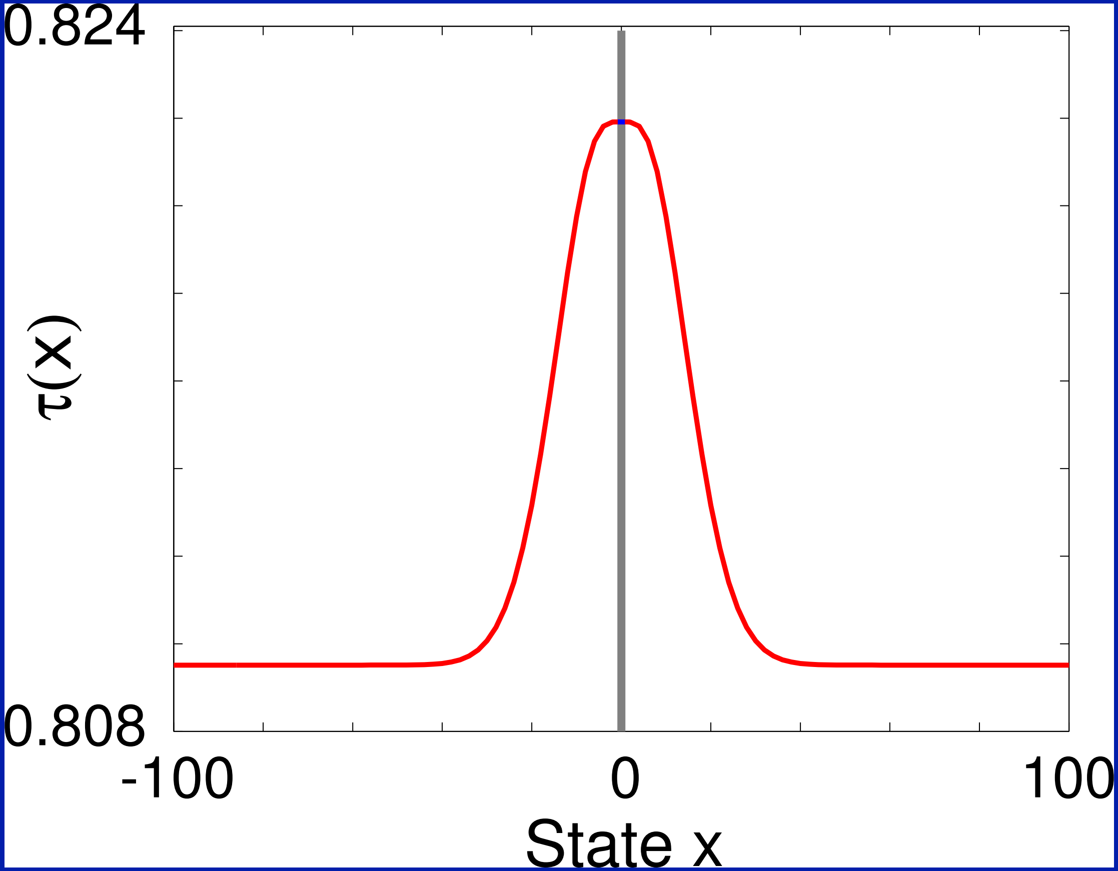
and $\tilde{A}(s,t)$: allowable values of α at (s,t)

With Adversarial Action

the solution for M=1 and general N

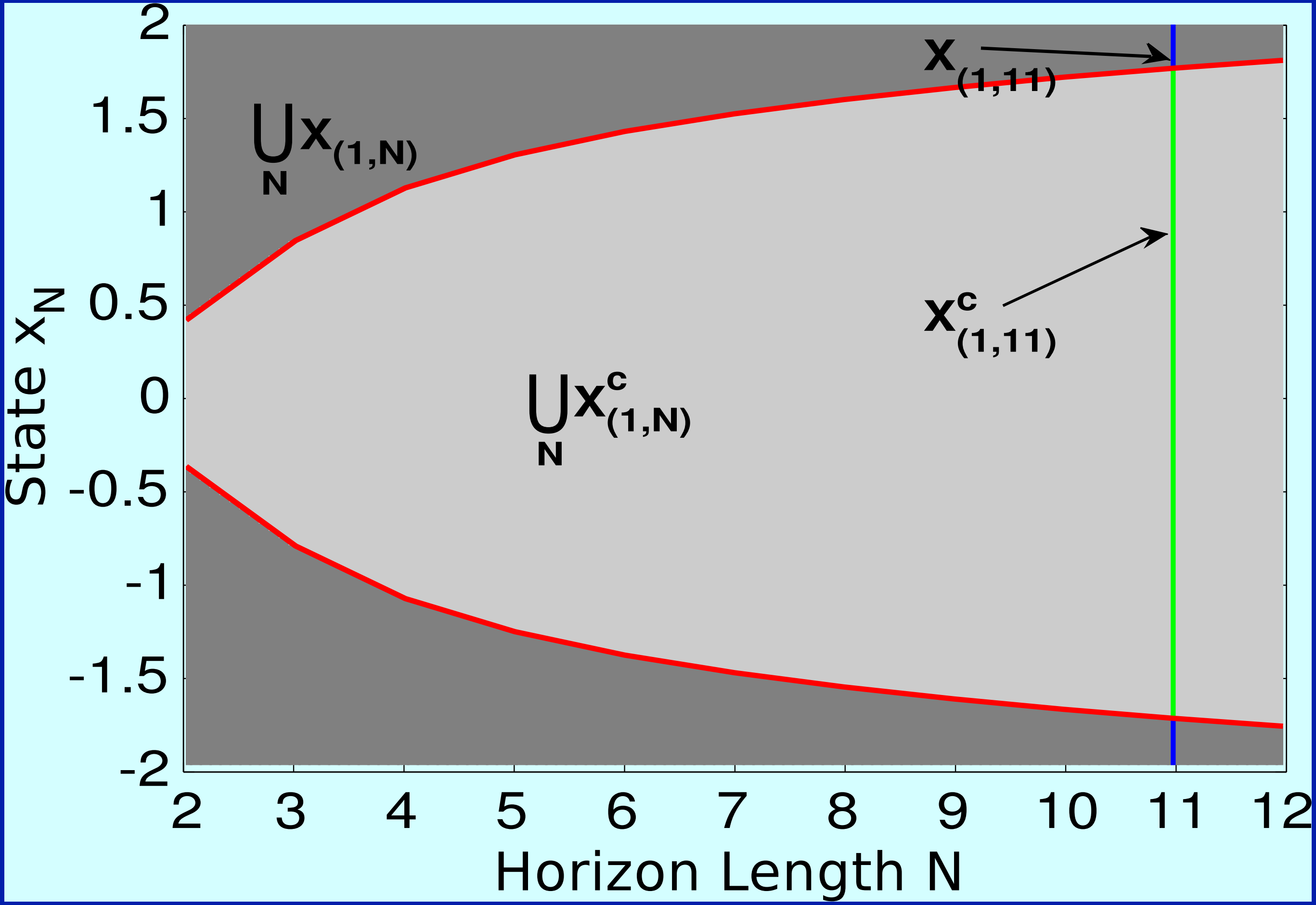
- There exists a saddle-point solution (*not in mixed strategies*)
- There exists a recursively computable threshold $\tau_{(s,t)}(x)$ such that
  - The jammer acts if $|x| - \tau_{(s,t)}(x) \geq 0$
  - The jammer does not act if $|x| - \tau_{(s,t)}(x) < 0$
- $V_{(s,t)}(x)$ admits two separate expressions depending on whether $|x| - \tau_{(s,t)}(x)$ is + or not
- Multi-dimensional case is qualitatively similar

Numerical study (A=2.5, σ_w = 1, N=3)
Plot of $\tau_{(1,3)}$ vs x

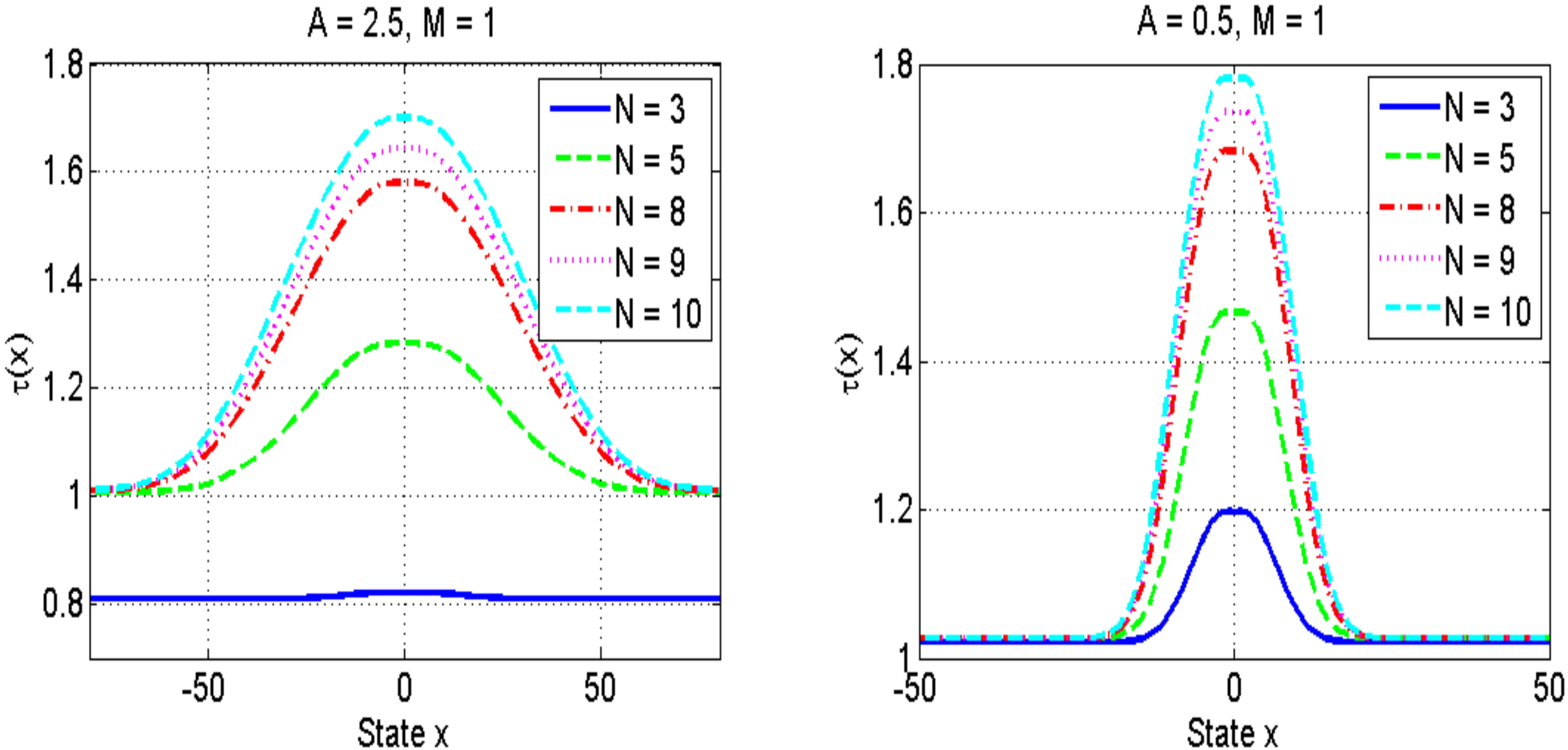


Jammer is active except in the narrow strip

Numerical study (A=2.5, σ_w = 1, general N)
Regions where jammer is active (dark)



Numerical study (σ_w = 1, N = 3, 5, 8, 9, 10)
Plot of $\tau_{(1,N)}$ vs x



Message to be taken

Opportunistic sensing, control, and decision making

Limitations on usage leads to event driven actions, where events are also controlled or caused by adversarial action

Returning to NCIS

- Limited memory leads to non-classical IS
- Having limits on frequency of actions of agents leads to non-classical IS (NCIC)
- Decentralization leads to NCIS
- Delays or disruptions in transmission leads to NCIS
- Private information also leads to NCIS; how much to reveal through actions?

Returning to NCIS

- Limited memory leads to non-classical IS
- Having limits on frequency of actions of

FERTILE GROUND

FOR RESEARCH

to NCIS

- Private information also leads to NCIS; how much to reveal through actions?

THANKS !