Optimal Stochastic and Networked Control Under Information Constraints

Serdar Yüksel

Queen’s University,
Department of Mathematics and Statistics

CAIMS 2013, Quebec City

Acknowledgements

Colleagues:
Tamer Başar (Univ. Illinois at Urbana-Champaign)
Giacomo Como (Lund Univ.)
Tamás Linder (Queen's Univ.)
Aditya Mahajan (McGill Univ.)
Nuno Martins (Univ. Maryland)
Sean Meyn (Univ. Florida)
Tobias Oechtering (KTH, Stockholm)
Mikael Skoglund (KTH, Stockholm)
Sekhar Tatikonda (Yale Univ.)

Graduate students:
Andrew Johnston, Naci Saldi, Nevroz Şen, Ali Zaidi

Colleagues, staff and students at Queen's.

Interaction of Information, Control and Probability

The *interaction between information and control* is a phenomenon that arises in every decision and control problem.

Any performance-driven controller requires information on the unknowns that affect the operation of the underlying system.

The uncertainty can typically be characterized through a (possibly incomplete) probability model.

The *quality* of the relevant information itself is typically affected by the choice of the control action in a closed-loop system.

Interaction of Information, Control and Probability

A collection of decision makers who wish to achieve *a common goal*, but who may have different (on-line) information variables, is said to be a *team*.

(This is in contrast to settings where the goals may not be aligned, as in *games*).

Networked Control refers to a team of decision makers (decentralized control system) in which the decision makers are connected through communication channels.

Thus, there may be a data link between the sensors (which collect information), the controllers (which make decisions), and the actuators (which execute the controller commands).

Networked Control

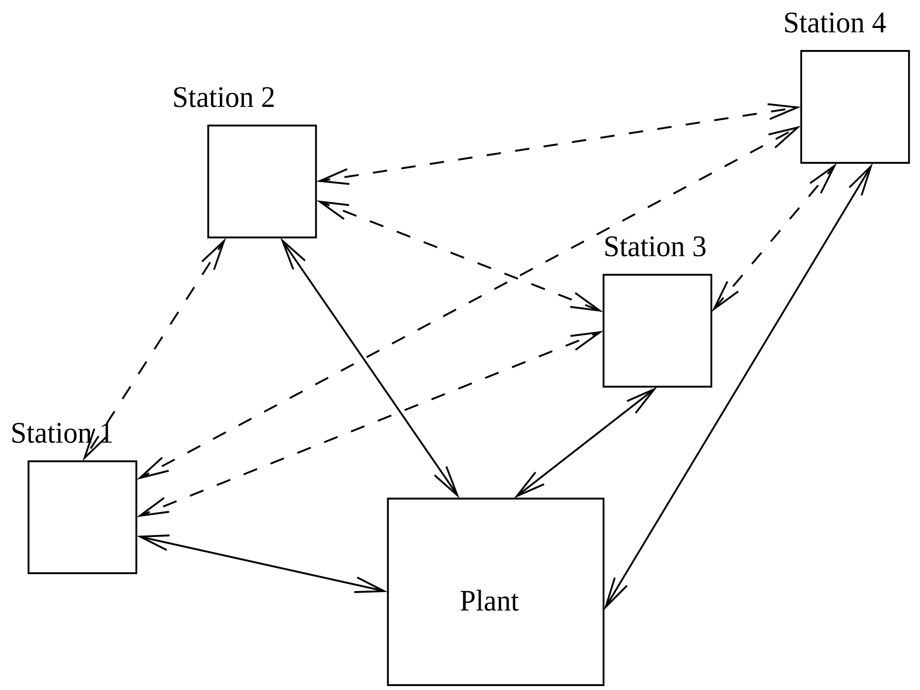


Figure 1: Decentralized control with rate-limited interactions.

Networked Control

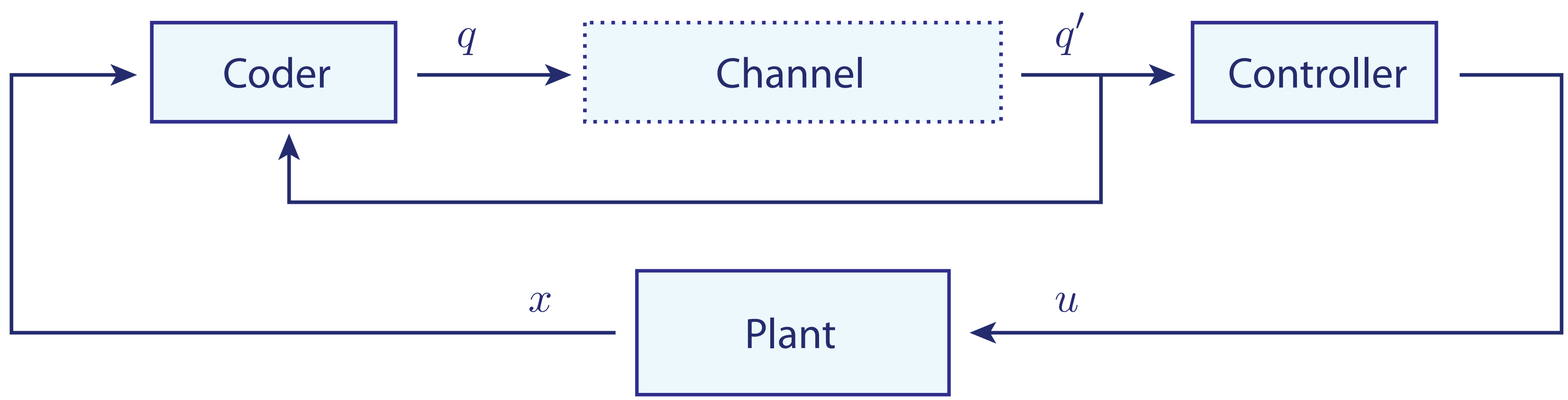


Figure 2: Another networked control system. Here, the coder and the controller are decision makers.

Networked Control and Information Structures

Such systems are becoming ubiquitous. Applications in:

Design of (smart) power grids,
Communications and networking systems
Automobile and inter-vehicle communications designs
Control of surveillance and rescue robot teams for access to hazardous environments
Space exploration and aircraft design
Control of economic systems
Theory of organizations in economics/business
among many other fields of applications which involve *decentralized decision making under uncertainty*.

In this tutorial talk, we will discuss stabilization, optimization and information structure related aspects of such systems, focusing first on single-channel systems.

Controlled Stochastic Systems

A controlled stochastic system has the following state / measurement equations:

$$x_{t+1} = f(x_t, u_t, w_t), \tag{1}$$

$$y_t = g(x_t, v_t) \tag{2}$$

A control policy γ is a sequence of control functions $\{\gamma_0, \gamma_1, \cdots\}$ each a causal function of the information vector

$$I_t = \{y_t; y_{[0,t-1]}, u_{[0,t-1]}\} \quad t \geq 1,$$

with control actions $u_t = \gamma_t(I_t)$.

In stochastic control, typically a measurement model is given and one looks for a control policy for optimization or stabilization of the controlled system.

Stochastic control theory is a rich, mature field with many applications.

Stochastic Control with Information Constraints

Here, (2) defines a channel, a stochastic kernel, mapping the state to measurements.

In networked control systems, the measurement channel itself and the information vector I_t are design variables: We can shape the observation / measurement channel, through encoding and decoding.

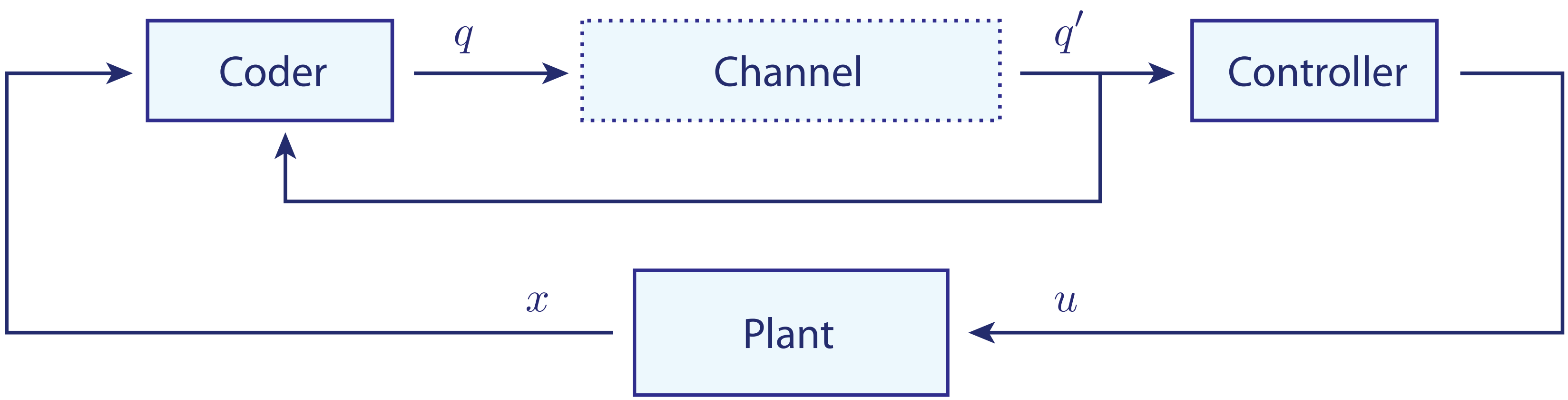


Figure 3: Encoding shapes the observation given the state. Joint design of coding and control is needed.

Design of Coder and Controllers

A *Coding Policy* Π is a sequence of functions $\{Q_t^c, t \geq 0\}$ such that the channel input, $q_t \in \mathcal{M}$, under Π is generated by a function of its local information

$$\mathcal{I}_t^e = \{x_{[0,t]}, q'_{[0,t-1]}\}$$

to the channel input alphabet $\mathcal{M}$. This is typically a finite set.

The channel maps q_t to q'_t.

A control policy γ is a set of functions $\{\gamma_t\}$ such that $u_t = \gamma_t(q'_{[0,t]})$.

Problem P1: Design of Information Channels for Stabilization

Given a system controlled over a channel, find the set of channels $\mathcal{Q}$ for which there exists a policy (both control and encoding), such that $\{x_t\}$ is *stochastically stable* in the following senses:

(i) recurrence / ergodicity / (asymptotic mean) stationarity,

and

(ii) existence of finite moments.

These will be specified further in this talk.

Problem P2: Design of Information Channels for Optimization

Given a controlled dynamical system, a channel, and a cost function c, the goal is to minimize

$$E^{\Pi,\gamma}\left[\sum_{t=0}^{T-1} c(x_t, u_t)\right], \tag{3}$$

over the set of all admissible coding and control policies Π, γ.

Find optimal coding and control policies; structural, existence and topological properties of such optimization problems.

One could also take the horizon to infinity, considering average/discounted cost.

Problem P1: Design of Information Channels for Stabilization

We will consider a linear Gaussian unstable system model (results are applicable to higher-order systems)

$$x_{t+1} = Ax_t + Bu_t + w_t,$$

It is assumed that there exists an unstable eigenvalue and (A, B) is controllable. For ease in presentation, assume A is diagonal.

Partially observed settings can also be considered.

This system is connected over a channel with a finite capacity to a controller.

Causal Coding for Control

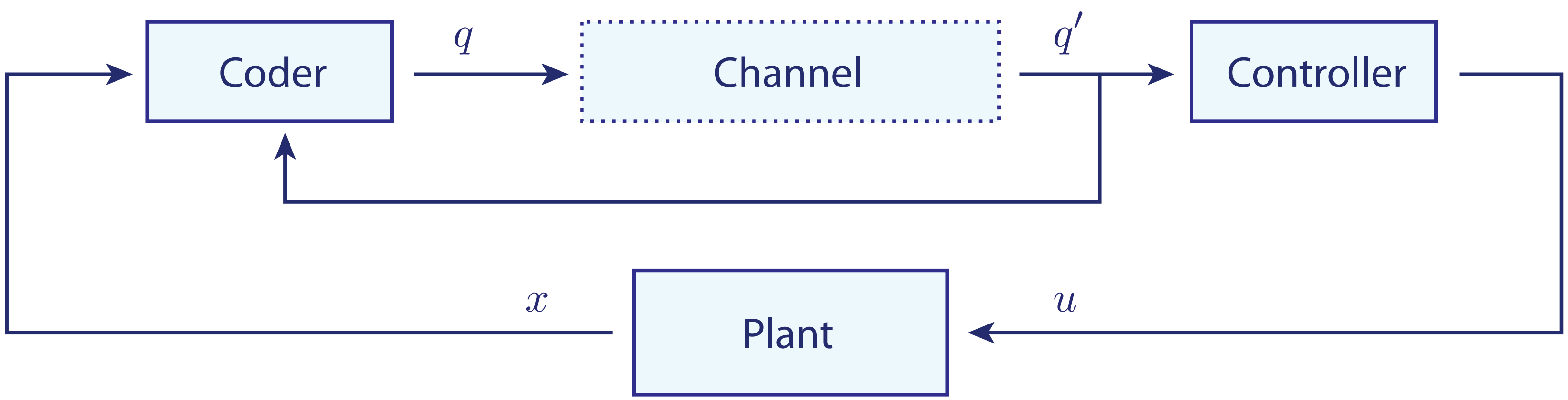


Figure 4: Control over a discrete noisy channel with feedback.

Channel Models and Channels used in Practice

A discrete noisy channel is a stochastic kernel such that for any $n \in \mathbb{N}$, an input sequence $q_{[0,n]}$ leads to an output $q'_{[0,n]}$ with probability $P(q'_{[0,n]}|q_{[0,n]})$. The channel is memoryless, if

$$P(q'_{[0,n]}|q_{[0,n]}) = \prod_{k=0}^{n} P(q'_k|q_k).$$

A **binary symmetric channel** is defined by the transition probabilities: $P(q' = 0|q = 1) = P(q' = 1|q = 0) = \epsilon$. **Binary noiseless** if $\epsilon = 0$.

A **binary erasure channel** is defined by the transition probabilities $P(q' = 0|q = 0) = P(q' = 1|q = 1) = 1 - \epsilon$ for some $\epsilon \in [0, 1]$, and an erasure symbol e such that $P(q' = e|q = 0) = P(q' = e|q = 1) = \epsilon$.

A **Gaussian channel** is one in which $q' = q + w$, where w is a Gaussian, independent noise variable.

Causal Coding for Control

The quantizer and the source coder policy is causal such that the channel input at time $t \geq 0$, q_t, is generated using the information:

$$I_t^s = \{x_{[0,t]}, q_{[0,t-1]}, q'_{[0,t-1]}\}$$

The quantizer outputs are transmitted through a channel, after being subjected to a channel encoder. The receiver has access to noisy versions of the quantizer/coder outputs for each time, which we denote by $q'_t \in \mathcal{M}'$.

The control policy at time t, also causal, only uses I_t^c, for $t \geq 0$:

$$I_t^c = \{q'_{[0,t]}\}$$

We will call such coding and control policies *admissible* policies.

Causal Coding for Control

We will consider the question: *When does a linear system driven by unbounded noise (such as Gaussian noise), controlled over a channel (possibly with memory) satisfy the following*:

- $\{x_t\}$ is asymptotically mean stationary and satisfies Birkhoff's sample path ergodic theorem (Stationarity).
- $\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} ||x_t||^2$ exists and is finite almost surely (Quadratic Stability).

Information Theory: Source Coding

Given an $\mathbb{X}$-valued source process $\{x_t\}$, and a distortion criterion ρ, a rate-distortion pair (R, D) is achievable if there exists a sequence of

-Encoders: $\mathcal{E}_n : \mathbb{X}^n \to \mathcal{M}(n)$ with $|\mathcal{M}(n)| \leq 2^{Rn}$

-Decoders: $\mathcal{D}_n : \mathcal{M}(n) \to \mathbb{X}^n$ such that $\hat{x}_t = \mathcal{D}_n(x_{[0,n-1]})$ with

$$D_n := \frac{1}{n} \sum_{i=0}^{n-1} \rho(x_n, \hat{x}_n) \leq D$$

inf $\{R : (R, D)$ is achievable$\}$ is the **Rate-distortion** function of a source.

Example: The rate-distortion function for a Gaussian source with variance σ_x^2 is given by $\frac{1}{2} \log_2(\sigma_x^2 / D)$ bits per source realization.

Information Theory: Channel Coding

Given a channel, a rate R can be reliably transmitted if there exists a sequence (R, ϵ_n), with $\epsilon_n \to 0$, such that for every n:

- There exists a set of messages $\mathcal{M}(n) := \{1, 2, 3 \ldots, M(n)\}$ such that $|\mathcal{M}(n)| \geq 2^{Rn}$

- A channel coder: $\mathcal{E}_n : \mathcal{M}(n) \to \mathcal{M}^n$ and a decoder: $\mathcal{D}_n : \mathcal{M}'^n \to \mathcal{M}(n)$, with average error probability

$$P_e := \frac{1}{|\mathcal{M}(n)|} \sum_{c \in \mathcal{M}(n)} P(\mathcal{D}(q'_{[0,n-1]}) \neq c | c) \leq \epsilon_n.$$

Given a channel, the supremum rate R can be achieved is called the **Capacity** of the channel.

Example: The capacity of an erasure channel is $(1 - \epsilon)$, where ϵ is the erasure probability.

Source Coding: Causal and Non-Causal

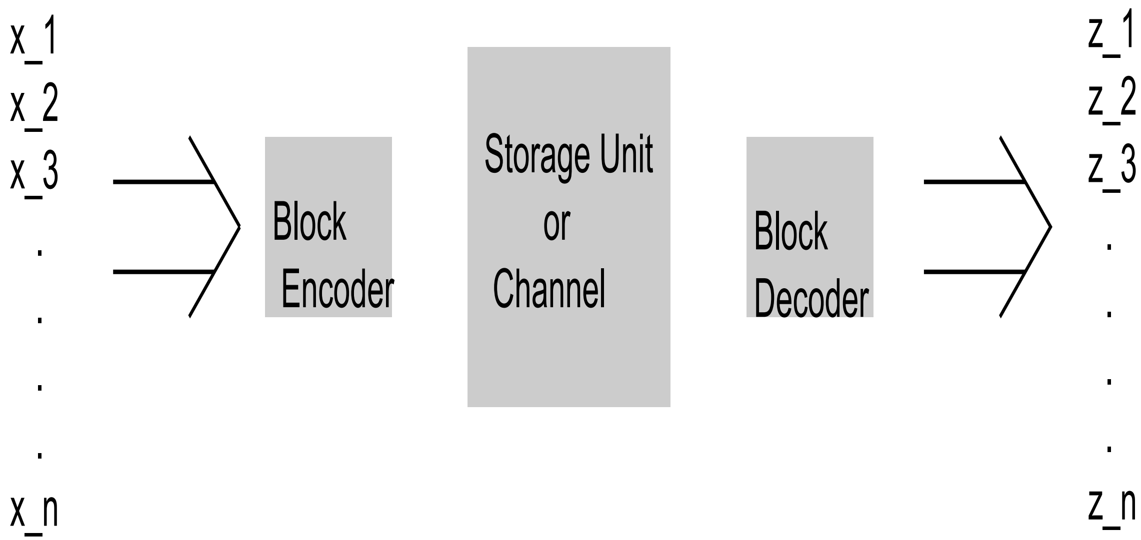


Figure 5: Information theoretic setup assumes non-causal block coders.

Are information-theoretic definitions useful for control problems?

The rate-distortion and capacity definitions are non-causal.

The encoder collects: $x_0, x_1, x_2, \ldots, x_n$, and generates $q_{[0,n]}$, as a block; then $\hat{x}_{[0,n]}$ is generated.

In control applications, the system cannot tolerate long delays:

$$x_0, q_0, q_0', u_0, x_1, q_1, \cdots$$

However, information theory suggests designs, and offers bounds on what is possible and not possible.

Literature Review: Information Theory for Unstable Processes

Consider the following Gaussian Auto-Regressive process:

$$x_t = -\sum_{k=1}^{m} a_k x_{t-k} + w_k,$$

where $\{w_k\}$ is an independent and identical, zero-mean, Gaussian random sequence with variance $E[w_1^2] = \sigma^2$.

If the roots of:

$$H(z) = 1 + \sum_{k=1}^{m} a_k z^{-k}$$

are inside the unit circle, the process is (asymptotically) stationary.

Literature Review: Information Theory for Unstable Processes

The rate distortion function (distortion being the normalized Euclidean error) is given parametrically by the following (Gray ('70), Hashimoto-Arimoto ('80), Berger ('70))

$$R(D_\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \max\left(\frac{1}{2}\log(\frac{1}{\theta g(w)}), 0\right) dw + \sum_{k=1}^{m} \frac{1}{2} \max\left(0, \log(|\rho_k|^2)\right), \quad (4)$$

with

$$D_\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \min(\theta, \frac{1}{g(w)}) dw,$$

where $\{\rho_k\}$ are the roots of the polynomial H.

Note that the encoding is non-causal.

Control Theory Literature and Causality Restrictions

Wong-Brockett'98, Nair-Evans'04, Tatikonda-Mitter'04 obtained that, for quadratic stability, an average rate of information transmission over a noiseless channel needed for stabilizability is at least

$$\sum_{k=1}^{m} \frac{1}{2} \max\left(0, \log(|\rho_k|^2)\right)$$

Contrasting with the Gray/Hashimito-Arimoto result, this shows that the rate term is not due to the causality restriction, but due to the uncertainty inherent in the source (the differential entropy rate).

There have also been other contributions by many researchers (Martins-Dahleh'08, Sahai-Mitter'06, Matveev-Savkin'08, Matveev'08 etc.) typically for bounded noise setups and under various notions of stability criteria such as finite second moments and stability in probability.

Causal Coding for Control: Noisy Channels

In **Problem P1**, the problem is to find, for the system

$$x_{t+1} = Ax_t + Bu_t + w_t,$$

the largest class of channels $\mathcal{Q}$, for which there exists a policy (both control and encoding), so that $\{x_t\}$ is stochastically stable:

When does an unstable linear system driven by unbounded noise, controlled over a channel (possibly with memory) is stochastically stabilizable in the following sense:

- The ergodic theorem applies, the state process is asymptotically mean stationary.

- The state process is quadratically stable, that is has finite average second moment which admits a limit.

The contribution is due the presence of unbounded noise, and the stronger criteria such as as stationarity, ergodicity and quadratic stability.

Asymptotic (Mean) Stationarity

Let $\mathbb{X} = \mathbb{R}^d$ and $\Sigma = \mathbb{X}^\infty$ denote the sequence space of all one-sided sequences drawn from $\mathbb{X}$. Thus, if $x \in \Sigma$ then $x = \{x_0, x_1, \ldots\}$ with $x_i \in \mathbb{X}$.

Let $X_n : \Sigma \to \mathbb{X}$ denote the coordinate function such that $X_n(x) = x_n$.

Let T denote the shift operation on Σ, that is $X_n(Tx) = x_{n+1}$. $Tx = \{x_1, x_2, \ldots\}$.

Definition .1. *A random process with measure P is $N-$stationary, (cyclo-stationary or periodically stationary with period N) if $P(T^N B) = P(B)$ for all $B \in \mathcal{B}(\Sigma)$. If $N = 1$, stationary.*

Definition .2. *A random process is $N-$ergodic if $B = T^N B$ implies that $P(B) \in \{0, 1\}$. If $N = 1$, it is ergodic.*

Stochastic Stability Notion: Asymptotic (Mean) Stationarity

Definition .3. *A process with a probability measure* $(\Omega, \mathcal{F}, P)$ *is asymptotically mean stationary (AMS) if there exists a probability measure* $\bar{P}$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} P(T^{-k}F) = \bar{P}(F),$$

for all events F*. Here* $\bar{P}$ *is the stationary mean of* P*.*

This property is equivalent to the applicability of the ergodic theorem.

Necessity and Sufficiency for Stationarity

Theorem .1. *[Y.'12,Y.'13]*

(i) For stability over a memoryless channel with any causal encoding and controller policy under the condition of the AMS property, the channel capacity must satisfy

$$C \geq \log_2(|A|) = \sum_{|\lambda_i| \geq 1} \log_2(|\lambda_i|).$$

(ii) If

$$C > \log_2(|A|) = \sum_{|\lambda_i| \geq 1} \log_2(|\lambda_i|),$$

there exist coding and control policies such that the state process is AMS.

This result also applies for a large class of channels with memory.

Necessity also holds for having $\limsup_{t \to \infty} E[|x_t|^2] < \infty$.

Skecth of Necessity through Information Theory

Mutual information satisfies

$$I(q'_t; q_{[0,t]}|q'_{[0,t-1]}) \geq I(x_t; q'_t|q'_{[0,t-1]})$$

$$\begin{aligned}
C &\geq \limsup_{T\to\infty}(1/T)\left(\sum_{t=1}^{T-1} I(x_t; q'_t|q'_{[0,t-1]})) + I(x_0; q_0)\right) \\
&\geq \limsup_{T\to\infty}(1/T)\left(\sum_{t=1}^{T-1}\left(h(x_t|q'_{[0,t-1]}) - h(x_t|q'_{[0,t]})\right) + I(x_0; q'_0)\right) \\
&= \log_2(|A|) - \liminf_{T\to\infty}\left((1/T)h(x_{T-1}|q'_{[0,T-1]})\right) \\
&\geq \sum_{|\lambda_i|>1} \log_2(|\lambda_i|),
\end{aligned}$$

Sufficiency

Let n be a given block length. We will consider a class of uniform quantizers, defined by two parameters, with bin size $\Delta > 0$, and an even number $K(n) \geq 2$:

$$Q^{\Delta}_{K(n)}(x) = \begin{cases} (k - \frac{1}{2}(K(n)+1))\Delta, & \text{if } x \in [(k-1-\frac{1}{2}K(n))\Delta, (k-\frac{1}{2}K(n))\Delta) \\ \mathcal{Z}, & \text{if } x \notin [-\frac{1}{2}K(n)\Delta, \frac{1}{2}K(n)\Delta), \end{cases}$$

where $\mathcal{Z}$ denotes the overflow symbol in the quantizer.

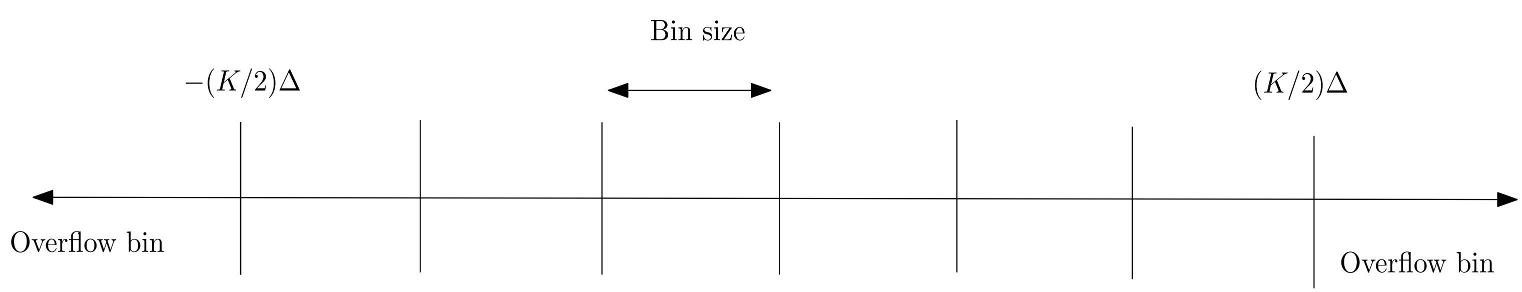


Figure 6: An adaptive uniform quantizer.

Sufficiency

Adaptive Zooming Quantizer: An approach known since early 70's (Goodman-Gersho (TCOM'74)). Also used in control literature recently extensively.

Zoom-out when the state has escaped the quantizer.

Zoom-in when the state is inside the quantizer's granular region.

Control is only applied when the state is inside the granular region. Inspired by this structure, the following drift criteria was obtained.

Stochastic Stability: Markov Chains

Let $\{x_t, t \geq 0\}$ be a Markov chain with state space $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$.

Definition .4. *For a Markov chain, a probability measure π is invariant on the Borel space $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ if*

$$\pi(D) = \int_{\mathbb{X}} P(x, D)\pi(dx), \quad D \in \mathcal{B}(\mathbb{X}) .$$

Existence of a unique invariant probability measure and positive Harris recurrence let the ergodic theorem hold:

$$\frac{1}{N}\sum_{t=0}^{N-1} f(x_t) \to \int \pi(dx) f(x),$$

for all integrable f under π a.s..

Stochastic Stability: Markov Chains under Random-Time State-Dependent Drift

The following characterizes stabilization when control is applied at random times. Let $\tau_z, z \geq 0$ with $\tau_0 = 0$ be such a sequence (of stopping times).

Theorem .2. *[Y. -Meyn (TAC'12)] Suppose that $\{x_t\}$ is a φ-irreducible Markov chain and $V : \mathbb{X} \to (0, \infty)$, $\delta : \mathbb{X} \to [1, \infty)$, $f : \mathbb{X} \to [1, \infty)$, a* small *set C, and a constant $b \in \mathbb{R}$, such that the following hold:*

$$E[V(x_{\tau_{z+1}}) \mid \mathcal{F}_{\tau_z}] \leq V(x_{\tau_z}) - \delta(x_{\tau_z}) + b1_{\{x_{\tau_z} \in C\}}$$

$$E\Big[\sum_{k=\tau_z}^{\tau_{z+1}-1} f(x_k) \mid \mathcal{F}_{\tau_z}\Big] \leq \delta(x_{\tau_z}), \qquad z \geq 0.$$

Then $\{x_t\}$ is positive Harris recurrent, and moreover $\pi(f) < \infty$, with π being the invariant distribution.

Stochastic Stability: Markov Chains

Corollary .1. *Suppose that $\{x_t\}$ is a φ-irreducible Markov chain. Suppose that there is a function $V : \mathbb{X} \to (0, \infty)$, a small set C, and a constant $b \in \mathbb{R}$, such that the following hold:*

$$E[V(x_{\tau_{z+1}}) \mid \mathcal{F}_{\tau_z}] \leq V(x_{\tau_z}) - 1 + b1_{\{x_{\tau_z} \in C\}}$$

$$\sup_{x \in \mathbb{X},\ z \geq 0} E[\tau_{z+1} - \tau_z \mid \mathcal{F}_{\tau_z} = x] < \infty.$$

Then $\{x_t\}$ is positive Harris recurrent (and there is a unique invariant measure, ergodic theorem holds).

Applying the Random-Time Drift Criterion for AMS

When the decoder output is the overflow signal, then the quantizer is zoomed-out.

Zoom-in when state is within the granular region of the quantizer. Define $h_t := \frac{x_t}{\Delta_t 2^{R'-1}}$. Apply control actions at these instants.

Define a sequence of stopping times (with n a block-length)

$$\tau_0 = 0, \quad \tau_{z+1} \quad = \quad \inf\{kn > \tau_z : |h_{kn}| \leq 1\}, \quad z, k \in \mathbb{Z}_+$$

For large Δ_{τ_z}, $\exists r > 1, M < \infty$: $P(\tau_{z+1} - \tau_z \geq kn | x_{\tau_z}, \Delta_{\tau_z}) \leq M r^{-kn}$.

Hence, by the random-time drift criterion, we can show that, there exist $\psi > 0$, $|G| < \infty$ such that

$$E[\log(\Delta^2_{\tau_{z+1}}) | \Delta_{\tau_z}, h_{\tau_z}] \leq \log(\Delta^2_{\tau_z}) - \psi + G 1_{\{|\Delta_{\tau_z}| \leq F\}} \tag{5}$$

Quadratic Stability - More Restrictive Conditions

For finite second moments, more restrictive conditions are needed. We take $V(x, \Delta) = \Delta^2$ as the Lyapunov function.

For erasure channels, we have a converse theorem, and an achievability theorem which are equal for a scalar source:

Theorem .3. *[Minero et. al.'09, Y.-Meyn'13] A necessary and sufficient condition for* $\limsup_{t\to\infty} E[|x_t|^2] < \infty$ *is:*

$$a^2\left(\epsilon + \frac{1-\epsilon}{2^{2R}}\right) < 1,$$

where ϵ is the erasure probability.

For more general channels, there is a gap between the converse and the achievability results. Details are omitted.

Problem P2: Design of Information Channels for Optimization

Consider the optimization problem where the controller has access to channel outputs where $Q(dy|x)$ is the channel.

We consider first the single-stage case

$$
\begin{aligned}
J(P,Q) &= \inf_{\Pi} E_P^{Q,\Pi}\big[c(x_0,u_0)\big] \\
&= \inf_{\gamma \in \mathcal{G}} \int_{\mathbb{X}\times\mathbb{Y}} c(x,\gamma(y))Q(dy|x)P(dx)
\end{aligned}
$$

in the channel Q, where $\mathcal{G}$ is the collection of all $\gamma : u = \gamma(y)$.

Problem P2: A Topology on Information Channels

Let $\mathcal{P}(\mathbb{X})$ denote the family of probability measures on $\mathbb{X}$.

Let $\{\mu_n,\ n \in \mathbb{N}\}$ be a sequence in $\mathcal{P}(\mathbb{R}^N)$. The sequence $\{\mu_n\}$ is said to converge to $\mu \in \mathcal{P}(\mathbb{R}^N)$ *weakly* if

$$\int_{\mathbb{R}^N} c(x)\mu_n(dx) \to \int_{\mathbb{R}^N} c(x)\mu(dx)$$

for every continuous and bounded $c : \mathbb{R}^N \to \mathbb{R}$.

Definition .5. [Convergence of Channels] (i) *A sequence of channels* $\{Q_n\}$ *converges to a channel* Q weakly at input P *if* $PQ_n \to PQ$ *weakly.*

(ii) *A sequence of channels* $\{Q_n\}$ *converges to a channel* Q *in* total variation at input P *if* $PQ_n \to PQ$ *in total variation, i.e., if* $\|PQ_n - PQ\|_{TV} \to 0$*.*

Continuity on the Space of Channels

Theorem .4. *[Y.-Linder'12]*

i) If $\mathbb{U}$ is compact and c is continuous, an optimal control policy exists for every channel.

ii) We do not have continuity in channels under weak convergence even for continuous cost functions.

iii) If the cost function is measurable and bounded, the optimal cost $J(P, Q)$ is continuous on the set of communication channels $\mathcal{Q}$ under under the topology of total variation.

This result will be useful to prove existence of optimal coding policies shortly.

Application: Quantizers as a class of channels

Definition .6. *An M-cell* vector quantizer, *q, is a (Borel) measurable mapping from $\mathbb{X} = \mathbb{R}^n$ to the finite set $\{1, 2, \ldots, M\}$, characterized by a measurable partition $\{B_1, B_2, \ldots, B_M\}$ such that $B_i = \{x : q(x) = i\}$ for $i = 1, \ldots, M$. The B_i are called the codecells (or bins) of q.*

A quantizer Q with cells $\{B_1, \ldots, B_M\}$, however, can also be characterized as a stochastic kernel Q on $\mathcal{B}(\mathbb{X} \times \{1, \ldots, M\})$ defined by

$$Q(i|x) = 1_{\{x \in B_i\}}, \quad i = 1, \ldots, M$$

so that $Q(x) = \sum_{i=1}^{M} Q(i|x)$.

In the analysis, we will restrict the quantizers to have convex codecells.

Existence of Optimal Quantization Policies

Let $\mathcal{Q}_C(M)$ denote the set of quantizers with convex codecells.

Theorem .5. *[Y.-Linder'12] The set $\mathcal{Q}_C(M)$ is compact under total variation at any input measure P that admits a density.*

Combined with continuity:

Theorem .6. *[Y.-Linder'12] Let P have a density and suppose the goal is to find the best quantizer Q with M cells minimizing $J(P,Q) = \inf_\gamma E_P^{Q,\gamma} \int c(x,u)$, for measurable and bounded c. Then an optimal quantizer exists.*

Multi-Stage Case: Dynamic Encoders

Consider a general controlled Markov system of the form, as before, with quantized information:

$$x_{t+1} \quad = \quad f(x_t, u_t, w_t),$$

Suppose that the goal is the minimization,

$$\inf_{\Pi^{comp}} \inf_{\gamma} E^{\Pi^{comp},\gamma}[\sum_{t=0}^{T-1} c(x_t, u_t)], \tag{6}$$

over all quantization policies Π^{comp}, and control policies γ.

Obtaining the **structure of optimal policies** is very important: The importance is both in obtaining analytical solutions and topological properties, but also to obtain computational complexity reductions.

A Structural Result on Optimal Quantization Policies

Theorem .7. *[Walrand-Varaiya'82, Y.'12] Under the objective given in (6), any* causal quantization policy *can be replaced, without any loss in performance, by one which only uses the conditional probability measure* $\pi_t(\cdot) = P(x_t \in \cdot | q_{[0,t-1]})$*, the state* x_t*, and the time information* t*, at time* t*.*

This can be expressed as a quantization policy which only uses $\{\pi_t, t\}$ *to generate a quantizer, where the quantizer uses* x_t *to generate the quantization output* q_t*.*

Thus, a quantizer can eliminate the past state data without any loss in optimality; $x_{[0,t-1]}$ can be eliminated!

This result also applies to partially observed Markov sources.

Stochastic Control Formulation

Let Π_W denote this class of optimal policies. Under such a policy, we obtain a non-linear filtering equation:

$$\pi_t(dx) = \frac{\int \pi_{t-1}(dx_{t-1}) P(q_{t-1}|\pi_{t-1}, x_{t-1}) P(dx|x_{t-1}, u_{t-1})}{\int \int \pi_{t-1}(dx_{t-1}) P(q_{t-1}|\pi_{t-1}, x_{t-1}) P(dx|x_{t-1}, u_{t-1})}.$$

Here, $P(q_{t-1}|\pi_{t-1}, x_{t-1})$ is determined by the quantizer policy.

The sequence of conditional measures and quantizers $\{\pi_t, Q_t\}$ form a controlled Markov process in $\mathcal{P}(\mathbb{R}^n) \times \mathcal{Q}$.

Stochastic Control Formulation

The cost to be optimized is:

$$\inf_{\gamma} J_{\pi_0}(\Pi^{comp}, \gamma, T) = E_{\pi_0}^{\Pi^{comp}}\left[\frac{1}{T}\sum_{t=0}^{T-1} c_0(\pi_t, Q_t)\right],$$

where

$$c_0(\pi_t, Q_t) = \sum_{i=1}^{M} \inf_{u\in\mathbb{U}} \int_{Q_t^{-1}(i)} \pi_t(dx) c_0(x, u).$$

Thus, traditional Markov Decision Processes formalism applies: Consider a system of the form: $x_{t+1} = f(x_t) + w_t$, where $\{w_t\}$ is i.i.d. Gaussian, action space $\mathbb{U}$ is compact and $c_0 : \mathbb{R}^n \times \mathbb{U} \to \mathbb{R}_+$ is (i) bounded and continuous or (ii) quadratic.

Existence

Let Π_W^C be the set of coding policies in Π_W with quantizers having convex codecells.

Theorem .8. *[Y.-Linder'12] For any $T \geq 1$, there exists a policy in Π_W^C such that*

$$\inf_{\Pi^{comp} \in \Pi_W^C} \inf_{\gamma} J_{\pi_0}(\Pi^{comp}, \gamma, T) \tag{7}$$

is achieved. Letting $J_T^T(\cdot) = 0$ and

$$J_0^T(\pi_0) := \min_{\Pi^{comp} \in \Pi_W^C, \gamma} J_{\pi_0}(\Pi^{comp}, \gamma, T),$$

the dynamic programming recursion holds.

$$TJ_t^T(\pi_t) = \min_{Q \in \mathcal{Q}_c} \Big(c(\pi_t, Q_t) + TE[J_{t+1}^T(\pi_{t+1}) | \pi_t, Q_t] \Big)$$

Linear Quadratic Gaussian Problem

Consider a Linear Quadratic Gaussian setup. Let $x_t \in \mathbb{R}^n$ and the evolution of the system be given by the following:

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + w_t, \\ y_t &= Cx_t + v_t, \end{aligned} \tag{8}$$

Here, $\{w_t, v_t\}$ are mutually independent, zero-mean i.i.d. Gaussian.

The goal is the computation of

$$\inf_{\Pi^{comp}} \inf_{\gamma} \frac{1}{T} E_{\nu_0}^{\Pi^{comp},\gamma}[\sum_{t=0}^{T-1} x_t' Q x_t + u_t' R u_t].$$

where, $Q \geq 0$, $R > 0$.

Linear Quadratic Gaussian Problem

Theorem .9. *[Y.'12]*

i) An optimal control policy is given by $u_t = L_t E[x_t | q_{[0,t]}]$, *where*

$$L_t = -(R + B' P_{t+1} B)^{-1} B' K_{t+1} A,$$

and

$$P_t = A_t' K_{t+1} B (R + B' K_{t+1} B)^{-1} B' K_{t+1} A,$$
$$K_t = A_t' K_{t+1} A_t - P_t + Q,$$

with $K_T = P_{T-1} = 0$.
ii) Coders in Π_W *with a predictive encoder structure is without any loss.*
iii) There exists an optimal coding policy in Π_W^C.

Separation of estimation error and control and lack of dual effect of control

Control performance is independent of optimal encoder performance [Borkar-Mitter'96, Nair et al'08, Y.'12].

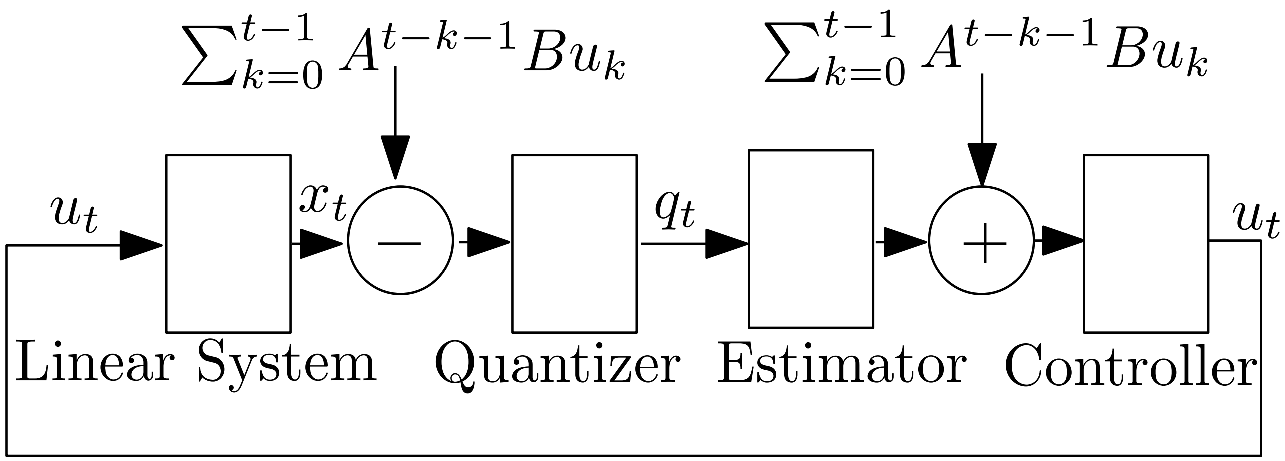


Figure 7: For the LQG problem, a predictive encoder is without loss.

Information Structures in Networked Control

Until this point, we discussed single-channel settings.

A systematic study of more general distributed systems require a careful classification of information structures.

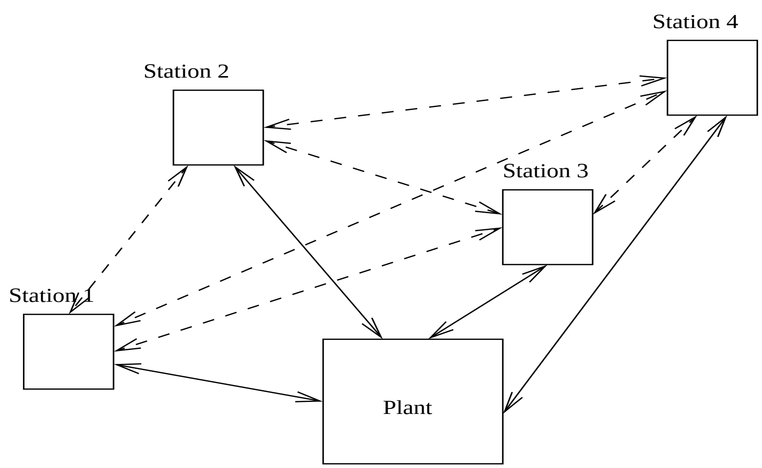


Figure 8: A decentralized networked control system.

Teams and Information Structures

Information structure in a networked control system is on which decision maker knows what information at any given time.

The information structure is a very important component in characterizing how difficult an optimization problem is.

Example: In single-decision makers, dynamic programming principle holds because information is *expanding*.

Example: In multi-decision maker settings, a dynamic programming approach is in general not possible: There is no *state* variable which is common knowledge! We can't apply the law of the *iterated expectation*.

Teams and Information Structures

A decentralized control system may either be *sequential* or *non-sequential*. In a sequential system, the decision makers (DMs) act according to an order that is specified before the system starts running.

If a decision maker's information does not depend on the action of another decision maker, the information structure is static.

There are some dynamic information structure problems which are equivalent to static ones (these are called quasi-classical).

Non-classical information structure arises if a Decision Maker (DM) i's action affects the information available to DM j, but DM j does not know what DM i knows.

Witsenhausen's Intrinsic Team Model

According to Witsenhausen, any (finite horizon) team problem can be characterized by a tuple $\big((\Omega, \mathcal{F}), N, \{U^i, i = 1, \ldots, N\}, \{\mathcal{J}^i, i = 1, \ldots, N\}\big)$.

- $(\Omega, \mathcal{F})$: The realization $\omega \in \Omega$ is called the *primitive variable* of the system. Ω denotes realizations and $\mathcal{F}$ is a set of events in Ω.
- N denotes the number of decision makers (DMs) in the system. Each DM takes only one action.
- $U^i, i = 1, \ldots, N\}$ is a collection of action spaces for each DM.
- $\{\mathcal{J}^i, i = 1, \ldots, N\}$ is a collection of sets in $\mathcal{F}$ and represents the *information* available to a DM to take an action. We can also show this with a measurement function η^i with range space I^i. The collection $\{\mathcal{J}^i, i = 1, \ldots, N\}$ is called the *information structure* of the system.
- A *control strategy (policy)*: $\{\gamma^i, i = 1, \ldots, N\}$ where $\gamma^i : \mathcal{J}^i \to U^i$.

Objectives and Policy Spaces

We will assume that we are given a probability measure P on $(\Omega, \mathcal{F})$ and a real-valued loss function ℓ on $(\Omega \times U^1 \times \cdots \times U^N) =: H$.

Any choice $\gamma = (\gamma^1, \ldots, \gamma^n)$ of the control strategy induces a probability measure P^γ on H. We define the performance $J(\gamma)$ of a strategy as the expected loss, i.e.,

$$J(\gamma) = E^\gamma[\ell(\omega, u^1, \ldots, u^N)]$$

where ω is the primitive variable and u^i is the control action of DM i.

Solutions to Convex Static Team Problems

Let

$$J(\gamma) := E[c(\omega, \gamma^1(\eta^1(\omega)), \ldots, \gamma^N(\eta^N(\omega)))].$$

A policy γ^* is *person-by-person-optimal* if for all k

$$J(\gamma^*) \le J(\gamma^{*1}, \ldots, \gamma^{*(k-1)}, \beta, \gamma^{*(k+1)}, \ldots), \qquad \beta \in \Gamma^k.$$

A policy γ^* is optimal if

$$J(\gamma^*) \le J(\gamma), \quad \gamma \in \Gamma = \Gamma^1 \times \Gamma^2 \times \cdots \times \Gamma^N$$

Person-by-person-optimality does not imply optimality, in general.

Theorem .10. [Radner'62] *For a static team problem with cost function $c(\omega, u^1, \cdots, u^N)$ which is (i) continuously differentiable and (ii) convex in the actions; a person-by-person-optimal strategy is globally optimal.*

Application: Static LQG Team Problems

Consider a two-controller system where x_1 is Gaussian and

$$x_2 = Ax_1 + B^1 u_1^1 + B^2 u_1^2 + w_1$$

$$y_1^1 = C^1 x_1 + v_1^1, \qquad y_1^2 = C^2 x_1 + v_1^2,$$

with w, v^1, v^2 zero-mean, i.i.d. disturbances. For $\rho_1, \rho_2 > 0$, let the goal be the minimization of

$$J(\gamma^1, \gamma^2) = E\left[||x_1||_2^2 + \rho_1 ||u_1^1||_2^2 + \rho_2 ||u_1^2||_2^2 + ||x_2||_2^2 \right] \tag{9}$$

over the control policies of the form: $u_t^i = \gamma_t^i(y_t^i), \quad i = 1, 2$.

Solution: By Radner's theorem, optimal policy is affine in measurements.
Remark: If quasi-classical, optimal solutions are still linear. If non-classical; they are not.

Non-Classical Settings and Lack of Convexity

Non-classical information structure arises if a Decision Maker (DM) i's action affects the information available to DM j, but DM j does not know what DM i knows.

For such a problem, DM i may want to *signal* her information to DM j through her control actions.

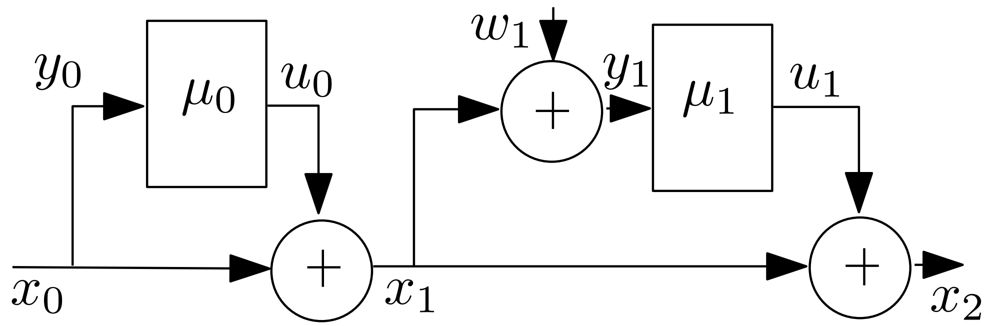
Note that the function

$$J(P,Q) = \inf_{\gamma \in \Gamma} \int c(x,\gamma(y))P(dx)Q(dy|x)$$

is concave in Q [Y.-Linder'10].

Hence, the *information structure design* or *channel design* is a *non-convex* problem, when there is active signaling.

Insight into Non-Classical Problems: Witsenhausen's Counterexample



$$y_0 = x_0, \qquad u_0 = \mu_0(y_0), \qquad x_1 = x_0 + u_0,$$
$$y_1 = x_1 + w_1, \qquad u_1 = \mu_1(y_1), \qquad x_2 = x_1 + u_1.$$

The goal is to minimize the expected performance index for some $k > 0$

$$Q_W(x, u_0, u_1) = k(u_0)^2 + x_2^2$$

Witsenhausen’s Counterexample

This is the celebrated Witsenhausen’s counterexample.

It is described by a linear system; all primitive variables are Gaussian.

Yet optimal team policy is *non-linear* [Witsenhausen’68].

Witsenhausen established that a solution exists and established that an optimal policy is *non-linear*.

The complete solution to this problem is not known. Properties of optimal solutions have been discussed in [Wu-Verdú’11].

Yet, if one replaces, in the Witsenhausen’s problem, the noise variables with binary variables, there does not exist an optimal team policy!

Optimal Dynamic Teams under Information Constraints

Efforts on existence, structure and approximations of optimal policies have ben studied.

Minimum amount of information to agree on certain *sufficient statistics* for optimal control has been studied with upper and lower bounds.

There is no *all-encompassing* systematic solution method in general for such problems:

Approaches for such problems typically follow from the construction of a controlled Markov state, information theory, reduction to static teams, expansion of information structures and so on [an extensive review is available in Y.-Başar (Springer'13)].

Concluding Remarks

For stabilization, there is a total order on the set of channels: For stationarity, Shannon theoretic capacity is the border for the converse and achievability.

For finite moments, the criteria we obtained are more stringent. Tight bounds for erasure channels have been obtained. Upper and lower bounds for more general channels are not equal.

Structural results and existence results for optimal coding and control policies, topological properties have been established.

A review of information structures has been presented: Information structure design in networked control is typically non-convex.

Even for LQG systems under non-classical information structures, linear policies are not optimal.

Future Directions

Distributed control under information constraints. Existence results, structural results and approximation techniques.

In multi-decision maker settings, a dynamic programming approach is in general not possible. In general, there is no *state* variable which is common knowledge: Sometimes, such a variable is available: Mean-field equilibrium [Huang et al.'07], Belief sharing pattern [Y.'09], information sharing [Nayyar-Mahajan-Teneketzis'11].

Connections with asymptotic agreement, Bayesian learning and consensus literature.

Learning/training methods and approximation techniques for computing optimal policies given structural and existence results.

Topological issues on mismatch in the beliefs/priors in optimal teams.

Value of information in *games*.