System Identification under Communication Constraints:
Complexity Aspects in Control-Communications Co-design

Le Yi Wang

Department of Electrical and Computer Engineering
Wayne State University
Once upon a time, in the land of cyber space, two groups of great people, control engineers and communications system specialists, were working diligently, happily, and separately in their own kingdoms, ....

In the kingdom of control systems, everyone was doing “feedback” which needed “information exchange”, or equivalently, “communication” or “communications”; But, most of the time, they could do it themselves ....
Years of Separation of Control and Communications

Mechanical Automatic Feedback Control Systems: 1788

Typical Feedback Control Systems with Sensors and Physical (mechanical, hydraulic, ...) or Wired Communications
Feedback System Design With Physical or Wired Communications

- Household Appliances
- Chemical Processes
- Power Systems
- Medical Devices
- Automotive Systems
- Communic'n Systems
- Computers
- Electric Machines
- Aerospace
- Weapon Systems

Command → Comparison → Error → Controller → Control Input → Controlled Plant → Output

Mechanical or Wired Data Transfer → Data Acquisition

Many Successful Applications of Traditional Feedback Control Systems
The Control Systems We Know Have the Key Foundations:

- **Additive Noise**
  Wiener filters, Kalman filters, predictive control, noise attenuation, tracking, system identification (equation noise, output noise, input noise), LQG, stochastic approximation, stochastic differential equations (diffusion terms), stochastic optimal control.

- **Uniform and Fast Sampling with ZOH**
  A LTI n-dimensional continuous-time system is sampled to a LTI n-dimensional discrete-time system.
  A LTI continuous-time system with a constant delay is sampled to a LTI (n+m)-dimensional discrete-time system.

- **Synchronous and Central Operation**
  In optimization, gradient calculation can be done and iterative updating is performed.
  In stability analysis, stabilization, robust control, optimal control, updating schemes are synchronous and central.

- **No Communication Resource Limitation**
  Signals are transported immediately with high resolution.
Now, look at more complicated networked systems
A Motivating Case: Smart Grids
Cyber Physical Systems

Sensor and Communications

Command to Physical Systems

Control and Decision

Subsystem 1

Subsystem 2

Subsystem n

Cyber Space

Physical Space

SMART GRID
A vision for the future — a network of integrated microgrids that can monitor and heal itself.
New Era of Networked Control Systems

New Information Age

Cloud Computing, IOT

Smart Transportation

System Biology

C³

Computing

Information

Physical Systems

Internet of Things

Communications

Smart City

Smart Grid

Intelligent Vehicles
Are networked control systems really new?
Demonstration at the center of Madison Square Garden

Nikola Tesla’s Wireless Controlled Boat 1898

Tesla: “the art of telautomatics”

Tel=Wireless Communications

Automatics= Control
Cybernetics

Norbert Wiener, Cybernetics or **Control and Communication** in the Animal and the **Machine**, MIT Press, 1948

Cybernetics is fundamentally **networked, feedback, learning/adaptation** systems with integrated control, communications, and computers

Thinkers of the Networked Systems and Information Era: 1940s-1950s
What factors are introduced by wireless communications that will impact control systems?
First, consider physical layer, dedicated, short distance, single user, wireless communications.
Simple Communications System Models
In Feedback Control Systems
(they are still used in many research papers)

They are within the traditional control toolboxes:
• Additive noise
• Fixed (maybe slowly time varying) time delays
• Some dynamics

Known problems: Communication People are Friends to Control!
But, how often do control systems enter the communications at the physical layer?

Open Systems Interconnection Model (OSI Model)

The Seven Layers of OSI

Control systems must comply with designated communications protocols

Example

DSRC (Dedicated short-range communications) are one-way or two-way short-range to medium-range wireless communication channels specifically designed for automotive use and a corresponding set of protocols and standards.
Scenarios from Communication Systems

Information → Quantization
Sampling → Data Compression

Data Packaging → Source Coding
Added Correction Bits

Receiver:
Error Detection and Correction
Decoding

Error! Resend!

Correct!
99.99% Accurate
No “additive noise” from the transmission

Correct Data Received For Processing
Complications that Communications Systems Introduce

- **Low Resolution Quantization:**
  Quantized Identification, State Estimation, Control, Optimization, Games

- **Irregular and random sampling times**
  LTI Continuous-Time Systems are sampled to Time-Varying Systems

- **Time varying and random delays**
  Time varying infinite dimensional stochastic systems

- **Non-uniform and distributed network operations:**
  Inconsistency in data (time-wise and contents). New asynchronous and distributed algorithms and their convergence and stability properties

- **Communication resource limitations:**
  Need complexity-based methodologies

All these constraints are new to state estimation, system identification, feedback control, stability, performance, robustness, and optimization.

All Hard problems: Communication People are Not Friends to Control!
Recent Advance
Dealing with Low Resolution Quantization
This is for low space complexity
Traditional State Estimation and System Identification

Many General Algorithms And Theoretical Results

Full Input Information

Full Output Measurement

Estimated State and System

Applications to Industrial, Biological Systems
New System Configurations

Signal recovering, state estimation, identification with quantized observations
Binary-Valued Sensors

Quantized Sensors

\[ y > C \quad \rightarrow \quad S = 00000000 \ldots \]

\[ y < C \quad \rightarrow \quad S = 11111111 \ldots \]

Not Sufficient Information to Know \( y \)!
Adding a “noise” \( d \), the output binary sequence \( s \) is now a random process.

Do we now have enough information to estimate \( \theta \)? YES!!!
Basic Mathematics Foundation

\[ y_k = \theta + d_k, \quad s_k = S(y_k), \]

binary-valued observation with threshold C

\[ \theta \] is to be estimated

\[ d_k \] is independent and identically distributed, \( d_k \sim F(x) \)

\[ s_k = 1 \iff y_k \leq C \iff d_k \leq C - \theta \]

\[ p = E[s_k] = P\{s_k = 1\} = P\{d_k \leq C - \theta\} = F(C - \theta) \]

Assume invertibility of \( F \):

\[ \theta = C - F^{-1}(p) \]

If \( p \) can be estimated, then \( \theta \) can also be estimated.

Empirical Measure (\( p \) estimation):

\[ \xi_N = \frac{1}{N} \sum_{k=1}^{N} s_k \]
Estimation of $\theta$: 
\[ \hat{\theta}_N = C - F^{-1}(\xi_N) \]
(remove the singular points $\xi_N = 0$ or $1$)

By the well-known Glivenko-Cantelli Theorem:
\[ \xi_N \to p, \text{ as } N \to \infty, \text{ with probability } 1 \]
\[ \hat{\theta}_N \to \theta, \text{ as } N \to \infty, \text{ w.p.1} \]

$\theta$ is asymptotically accurately estimated!!

- Noise is useful. In communications, “statistical quantization” is used to smooth jumps. In neural systems, they collectively transfer information in this way.
- This trades space complexity (less bits per sample by quantization, information loss) with time complexity (use more samples to recover information)
- This introduces nonlinearity into identification of a linear system
- Powerful tools of stochastic analysis can now be used.
Extensive Research Effort in the Past 10-15 years …
Extension to Finite Impulse Response Models

Extension to Rational Models

Joint Estimation of Systems and Unknown Noise Distributions

Extension to Nonlinear Wiener and Hammerstein Models

Identification Input Design

Quantized Identification under Data Rate Constraints

Error Probability with Large Deviation Principles

System Complexity Analysis

Extension to General Nonlinear Systems

Applications to Monitoring, Diagnosis, Networked Control, Games, ....
But, is this quantization-sampling framework really an efficient way to transport information and use communication resources?

There are much better ways to convey information if control and communication people work together.

Think how secret agents convey information: They code a lot of information into a very simple code.
Recent Advance

Sampling: Maximum Information from a Bit

A complexity issue, again
Communication Impact and Irregular Sampling

Issues

• Signals contains information
• Communications of signals need resource (power and bandwidth)

Question on Information vs Complexity

How much information about a signal that can be obtained after sampling and quantization?

That depends on sampling and quantization schemes.
Sampling and Quantization Schemes

Clocks at the sending and receiving sites are synchronized, so the sampling time itself does not need to be transmitted.

1. Sampling time is uniformly spaced.
2. Sampled values are known only within the quantization levels.
3. Fast sampling may be wasted.

This is not a desirable sampling scheme.
1. Sampling time is irregular.
2. Unnecessary samples are avoided. Communication resources are saved.
3. Sampled values are accurate (if no measurement noises are considered).
4. No guarantee on how many sampled points are generated.

More efficient sampling scheme, but issues with state estimation and control need to be resolved: It may not generate any sampling points for a long time.
1. Sampling time is irregular.
2. Number of sampled points per unit time interval is guaranteed by the carrier frequency.
3. By using synchronized clocks and the known carrier at both sending and receiving sites, communications will only be binary bits.
4. Sampled values are accurate, with an additive noise due to clock synchronization errors.

Effective sampling scheme, control of sampling density, issues of irregular sampling with state estimation and control need to be resolved.
Much past effort on irregular sampling and switching/hybrid systems:

- J. Raisch (1994)
- B.B. Astrom and B. Bernhardsson (1999)
- F. Ding, L. Qiu, and T.W. Chen (2009)
- N. Persson and F. Gustafsson (2001)
- C.E. de Souza, R.M., Palhares, and P.L.D. Peres (2001)
- M. Babaali and M. Egerstedt (2003, 2005)

............... (much more)

These efforts are control focused: What can we do if signals arrive irregularly?
But, what do we really need from communication people?

A complexity problem, again.
Example: State Estimation for Linear Systems

\[
\dot{x}(t) = Ax(t) + Bu(t) \\
y(t) = Cx(t)
\]

Results of irregular sampling:
A sequence of switching times and noise-corrupted sampled values

\[
t_1, t_2, \ldots, t_N, \ldots \text{ and } z(t_k) = y(t_k) + d_k
\]

State Estimation Problem

Estimate \( x(0) \) based on the noise-corrupted output observations with irregular sampling times.
Main Issues

\[ t_1, t_2, \ldots, t_N, \ldots \text{ and } z(t_k) = y(t_k) + d_k \]

**Question:** Does this time sequence provide sufficient information?

Observability

\[ d_k = 0. \, N \geq n \]

Will \( t_1, t_2, \ldots, t_N, \text{ and } z(t_k) = y(t_k) \) be sufficient for estimating \( x(0) \)?

In general, the answer is No!
Observation Relationships

At the switching times \( t_i, i = 1, \ldots, N \)

\[
y(t_i) + d_i = z(t_i) = C \, e^{A(t_i-t_N)} \, x(t_N) + C \int_{t_i}^{t_N} e^{A(t_i-\tau)} \, Bu(\tau) \, d\tau
\]

\[
C \, e^{A(t_i-t_N)} \, x(t_N) = y(t_i) - v(t_i, t_N) + d_i
\]

\[
\Phi_N \, x(t_N) = \Gamma_N - V_N + D_N
\]

\[
\Phi_N = \begin{bmatrix}
Ce^{A(t_i-t_N)} \\
\vdots \\
Ce^{A(t_{N-1}-t_N)} \\
C
\end{bmatrix}, \Gamma_N = \begin{bmatrix}
y(t_1) \\
\vdots \\
y(t_{N-1}) \\
y(t_N)
\end{bmatrix}, V_N = \begin{bmatrix}
v(t_1, t_N) \\
\vdots \\
v(t_{N-1}, t_N) \\
0
\end{bmatrix}, D_N = \begin{bmatrix}
d_1 \\
\vdots \\
d_{N-1} \\
d_N
\end{bmatrix}
\]

Observability \( \iff \Phi_N \) is full rank
General Sampling Theorem for Systems

Let the eigenvalues of $A$ be $\lambda_1, \ldots, \lambda_n$. \( \omega = \max_{i} \left| \text{Im}(\lambda_i) \right| \)

For $T > 0$, define \( \mu_T = 2(n-1) + \frac{T \omega}{\pi} \)

**Theorem**

Suppose the system is observable, $0 \leq t_i \leq T, i = 1, \ldots, N$ and $N > \mu_T$. Then $\Phi_N$ is full rank.

Asymptotically, \( \frac{\mu_T}{T} \approx \frac{\omega}{\pi} = \mu \), Characteristic Frequency Bandwidth of the System

If (sampling density) $N/T > \mu$, the state information on the system can be completely recovered from its sampled values.
A Comparison to Shannon’s Sampling Theorem

Shannon’s Sampling Theorem

For signal reconstruction
2. Periodic sampling
3. Non-causal signal reconstruction
4. Critical complexity relationship

\[ \frac{N}{T} > \frac{\omega_B}{\pi} \]

Nyquist frequency of periodic sampling vs. signal bandwidth
1. For linear systems
2. Irregular sampling
3. Causal state reconstruction
4. Critical complexity relationship

Let the eigenvalues of $A$ be $\lambda_1, \ldots, \lambda_n$. $\omega_B = \max_i |\text{Im}(\lambda_i)|$

$$\frac{N}{T} > \frac{2(n-1)}{T} + \frac{\omega_B}{\pi}$$

**Sampling density vs. system characteristic bandwidth**

Note:
1. For linear finite dimensional systems, its outputs are always band unlimited. So, Shannon’s sampling theorem cannot be applied here. Also, Shannon’s theorem is for uniform sampling only.
2. We do have more information: The modes of the signal are known from the $A$ matrix.
This complexity relationship shows up by itself in:

**Extension to Convergence, Convergence Rates, Estimation Accuracy**

**Extension to Joint Estimation of State and Events in Hybrid Systems**

**Extension to System Identification**

**Extension to Controllability**
Recent Advance
Decision and Complexity Based System Identification
Resource Allocation in Networks:
A complexity problem, again

Traditional System Identification

Example:

\[ y_k = \phi_k^T \theta + d_k, \quad k = 1, \ldots, N \]

Algorithms:

\[ \hat{\theta}_N = \left( \Phi_N^T \Phi_N \right)^{-1} \Phi_N^T Y_N, \quad \Phi_N = \begin{bmatrix} \phi_1^T \\ \vdots \\ \phi_N^T \end{bmatrix}, \quad Y_N = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \]

Goals:

Convergence: \( \theta_N \to \theta \), w.p.1, \( N \to \infty \)

Convergence Rate: \( NE(\theta_N - \theta)^2 \to \Sigma, \ N \to \infty \)

\[ \sqrt{\frac{N}{\log \log N}} (\theta_N - \theta) \to O(1), \ \text{w.p.1,} \ N \to \infty \]

Asymptotic Normality: \( \sqrt{N} (\theta_N - \theta) \to \mathcal{N}(0, \Sigma), \ N \to \infty \)

Asymptotic Efficiency: Achieve the CR Lower Bound, \( N \to \infty \)

Common Issue:

\( N \to \infty \)

Complexity Resource Money

Problem: We cannot spend the money we do not have.
Decision-Based Identification

First Question: How accurate should the estimates be?
That depends on what the “decisions” you must make.
Decisions: control, monitoring, diagnosis, prediction, coordination, etc.

Example 1: Robust Feedback Controller
If the controller is very robust, then identification accuracy can be reduced.
But if the plant is close to the boundary of the robust region, identification accuracy needs to be enhanced for controller adaptation.

Example 2: Patient Vital Sign Monitoring
If the patient is healthy, then identification accuracy can be reduced.
But if a patient is sick with blood pressures near the hypertension thresholds, a closer monitoring is needed for patient safety.

Example 3: Battery State-of-Health Estimation
Suppose that an EV battery must retire when its maximum capacity is reduced below 72-76% of the original rated capacity.
If the capacity remains above 85%, its estimation accuracy can be reduced.
But if a battery’s capacity declines below 80%, more accurate estimate is demanded.
Finite Resource Allocation Examples

$m$ users are sharing a communication channel. All users update their estimates of their own parameters once every $T$ seconds.

**Resource Allocation Strategy 1**

All users are assigned the equal number of time slots during the updating time interval $T$.

*Is this a good strategy?*
What is Decision and Complexity Based Identification?

Decision Reliability
(Probability) 1-\( \alpha \)

Decision Error
(Probability) \( \alpha \)

\( \theta_N = \theta^* + e_N \quad \Leftrightarrow \quad \text{error distribution depends on } N \)

For a given decision reliability 0 < \( \alpha \) < 1 (typically close to 0), the minimum resource assignment \( N^*(\theta^*) \) is defined as

\[ N^*(\theta^*) = \min\{N : \text{such that } \Pr_{e_N} \{\theta_N \in S \mid \theta^* \in M \} \geq 1 - \alpha\} \]
Improved decision reliability by using more data: Larger resource

Probability Characterization of Estimates

Use the **Central Limit Theorem and Large Deviations Principles** to obtain asymptotically accurate values for the probabilities of missed diagnosis and to guide decision algorithms to improve it.

Improved decision reliability by using more data: Larger resource

The tail areas are decision errors

Blood Pressure Monitoring Example

- **P***: normal
- **C***: hypertension
- **C**: systolic blood pressure **P**
Equal Resource Allocation (Population Based)  
• The overall reliability is determined by the worst case  
• Many resources are wasted  

Large Reliability Variations (Individual Reliability)  

Uniform Reliability  
Dynamic Resource Allocation (Individualized and Need Algorithms)  
• The uniform reliability  
• Resources are saved
Set identification looks like a hypothesis testing problem.

The main difference: We want to obtain the minimum resource assignment dynamically and individually.

Main Complication: The true parameter is unknown, and hence the optimal resource allocation cannot be done off-line.

Can we obtain the minimum resource assignment for the individual system in real time?
Adaptive Resource Allocation

New Estimation Problem for finding the optimal $N^*$:

1. Estimation Algorithms for $N_k$

2. Convergence: $N_k \to N^*$, w.p.1, $k \to \infty$

3. Convergence Rate: $kE(N_k \to N^*)^2 \to \Sigma$, $k \to \infty$

$$\sqrt{\frac{k}{\log \log k}} (N_k \to N^*) \to O(1), \text{ w.p.1., } k \to \infty$$

4. Asymptotic Normality: $\sqrt{k}(N_k \to N^*) \to \mathcal{N}(0, \Sigma)$, $k \to \infty$

5. Asymptotic Efficiency: Achieve the CR Lower Bound, $k \to \infty$

Good News: All these properties have been established for individual parameters under Gaussian i.i.d. cases

Good/Bad News: This is really a first baby step in this direction.
Summary

Cyber-Physical Integration
And
Control-Communication Co-Design
Our 3D Complex World of Networked Systems

**Data**
- Signal Quantization
- Data Compression
- Data with Limited Information

**Space**
- Data Locations Distributed physically and in cyber space
- Distributed Computation
- Distributed Decision but for a Global Mission

**Time**
- Irregular and Random Sampling
- Time-varying and Random Delay
- Asynchronous Operation
Kingdom of Control Systems
- Quantization
  - Sampling
  - Coding
  - Communication Power
  - Communication Bandwidth
  - Interference Avoidance
  - Priority

Kingdom of Communications
- Stability
- Performance
- Robustness
- Decision Accuracy
- Diagnosis
- Monitoring

Complexity is a Central Theme
- Latency
- PDR
- Throughput
- Stability
- Performance
- Robustness
- Decision Accuracy
- Diagnosis
- Monitoring

Managing resource allocations for
- Latency
- PDR
- Throughput

Request for controlled
- Delays
- Data rate
- Network topology

Co-Design:
Integrated Management of Resources from the Two Great Kingdoms
Then, these two groups of great people, control engineers and communications system specialists, are working diligently, happily, and collaboratively in their combined kingdoms, ever after ....
Thank You!