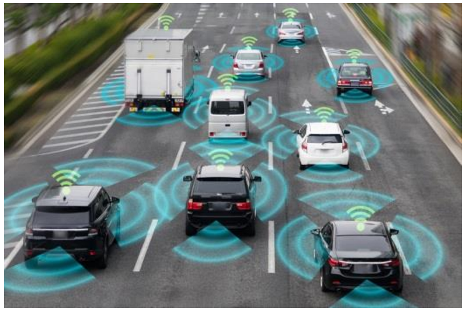
WAYNE STATE UNIVERSITY

# Integration of Control, Communications, and Computing in Networked Systems

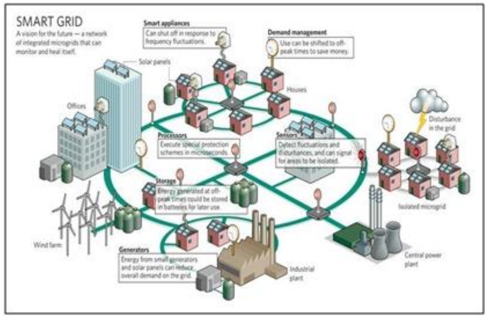
**Le Yi Wang**
**Department of Electrical and Computer Engineering**
**Wayne State University, USA**

China, July 2019

# New Systems are Increasingly Interconnected



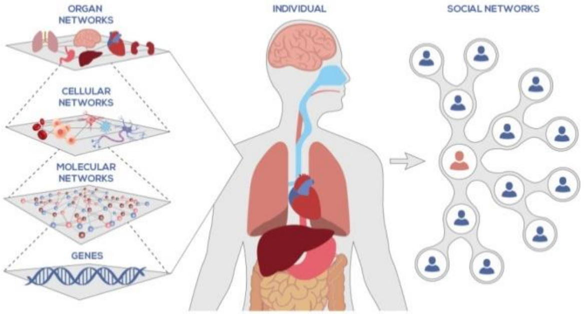
Connected Vehicles



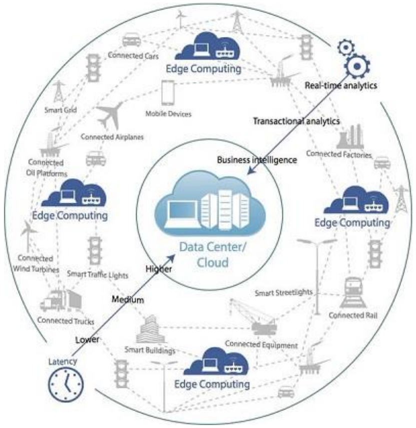
Smart Grids

Social Networks

Smart City



System Biology



Cloud and Edge Computing Network

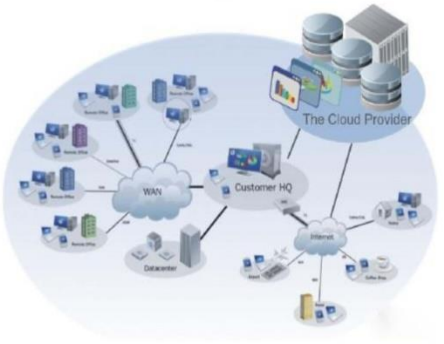
# New Information Age

$C^3$

Computing

Control

Communications

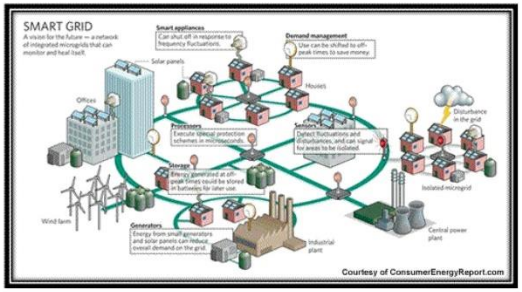


**IoT and Cloud**

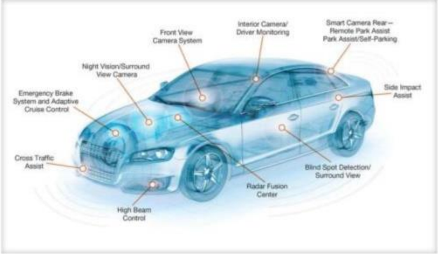
**Intelligent Transportation**

**Biological Systems**

**Smart City**



**Smart Grid**



**Intelligent Vehicles**

# Challenges and Opportunities for Paradigm Shifting in Control Systems

# Wiener's Integration of Control, Communication, and Computing 1935-1948

Von Newman Architecture ← **Digital Computers**

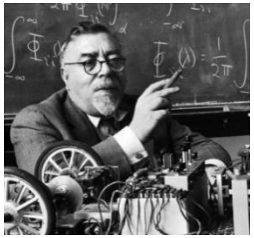
John Von Newman

**Communication Theory** → Shannon Communication and Information Theory

Claude Shannon

**Analog Computers**

Yuk Wing Lee (李郁荣)



**Cybernetics**

Feedback

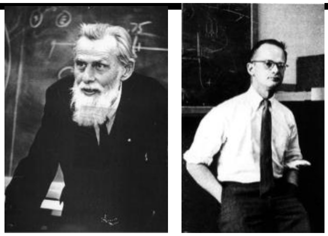
- Error Correction
- Adaptation to Environment
- Information Learning

↓

Control Theory

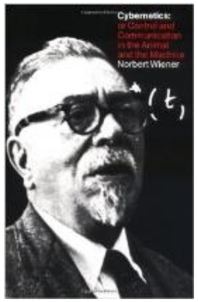
**Brain Functions**

Warren McCulloch
Walter Pitts

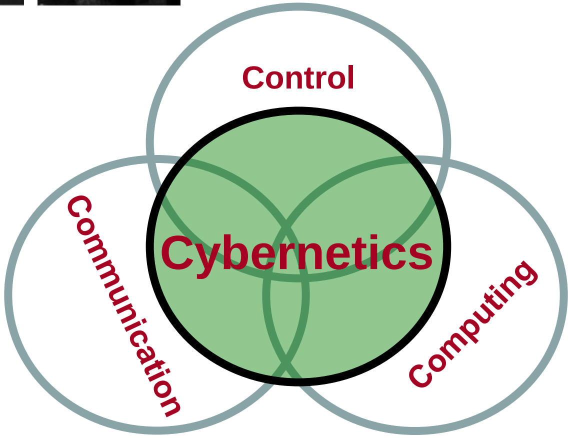


→

- McCulloch-Pitts Neuron Model
- Artificial Neural Networks
- Artificial Intelligence



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



# Fundamental Feedback Ability and Limitations

**Norbert Wiener**

**Yuk Wing Lee**

**George Zames**

G. Zames, "**Nonlinear operators** for systems analysis," Sc.D. dissertation, MIT, Cambridge, 1960

G. Zames, On **the input-output stability** of time-varying nonlinear feedback systems: Part I and Part II: IEEE TAC, Vol. 11-3, 1966

**1960s Foundation for Feedback Robustness, Small Gain, Sector Criteria, Passivity**

**Late 1970s: H-infinity Theory**

G. Zames, **Feedback** and **optimal sensitivity**: Model reference transformations, **multiplicative semi-norms** and approximate inverses," IEEE TAC, Vol. 26, Apr. 1981.

**Uncertainty models must be consistent with input-output behavior.** → **Return to transfer functions and operators (from state space models)**

**Norms must be consistent with subsystem connections: Multiplicative Norms** → **L2/H2 norms to H-infinity norm: System connections (networked systems)**

**Quantified feedback ability: Optimal robustness in stability and performance** → **Critical for complexity-based theory on connected systems**

G. Zames, On the **metric complexity** of causal linear systems, ε-entropy and ε-dimension for continuous-time systems, IEEE TAC, 1979.

**Solved Problems: Robust Stability, Nominal Sensitivity Minimization, State-Space Algorithms**

**More Difficult (but Important for Complexity Analysis) Problem: Robust Performance**

J. Owen and G. Zames, "Duality theory of MIMO **robust disturbance rejection**, IEEE TAC, May1993.
Numerical algorithms: Stephen Boyd (iterative convex optimization)

# Late 1980s-1997: Beyond H-infinity

**Building “learning” ability to H-infinity**

**Slowly time varying systems and H-infinity Adaptation**

George Zames and Le Yi Wang, Local-global double algebras for slow **H-infinity adaptation**, Part I; **IEEE TAC** 1991.
Le Yi Wang and George Zames, Local-global double algebras for slow H-infinity adaptation, Part II, **IEEE TAC** 1991.

**Complexity-based learning from data**

**Identification: Model Complexity (Kolmogorov n-widths), Time Complexity (Gelfand n-widths. Identification n-widths), Uncertainty Principles**

George Zames, Lin Lin and Le Yi Wang, **Fast identification n-width** and **uncertainty principle** for LTI and slowly time-varyng systems, **IEEE TAC**, Vol. 39, pp. 1827-1837, 1994.

- Lin Lin, Le Yi Wang, and George Zames, **Time complexity and model complexity** of fast identification of continuous-time LTI systems, **IEEE TAC**. 1999.
- L.Y. Wang, **Persistent identification** of time-varying systems**, IEEE TAC**, 1997.

**Feedback Organization**

**Decentralized Feedback for Uncertainty Reduction; Coordinating Feedback for Robust Performance**

George Zames, Towards a general complexity-based theory of identification and adaptation, LNCIS 222, A.S. Morse, Ed., Springer, 1997

L.Y. Wang and L. Lin, Information-based complexity of uncertainty sets in feedback control, **IEEE TAC**, Vol. 46, No. 4, pp. 519-533, April 2001.

**Combining Deterministic and Stochastic Frameworks**

**Private Communications with Le Yi Wang (unfinished work)**

Le Yi Wang and George Yin, Persistent identification of systems with **unmodelled dynamics** and **exogenous disturbances**, **IEEE TAC**, Vol. 45-7, 2000.

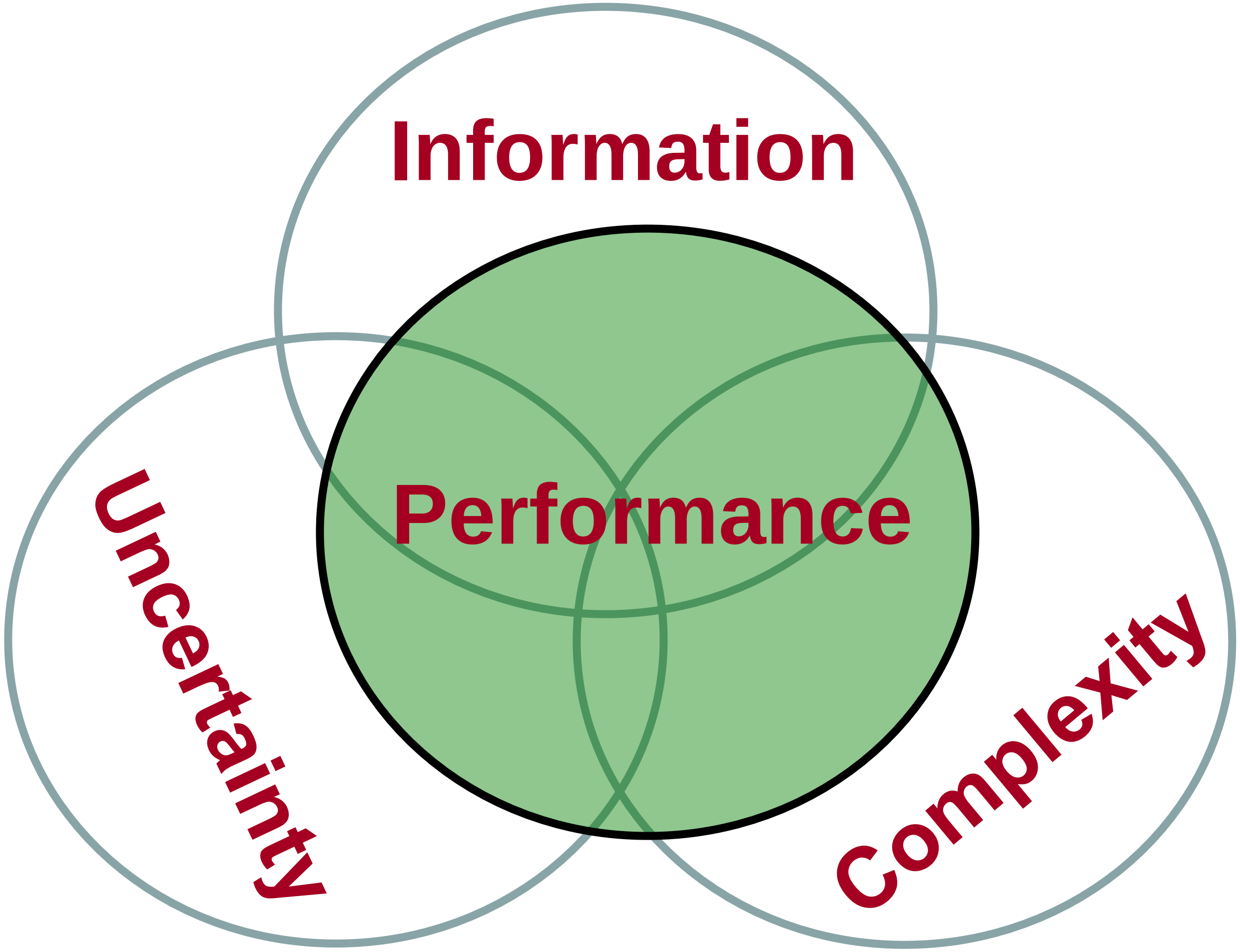
**– – – – -  Died 1997  – – – – -**

**My Best Guess:**

**Fundamental Complexity-Based Feedback (Robust, Adaptation, Learning) Framework for Cybernetics (Networked Systems)?**

# Uncertainty and Complexity
# In
# Networked Systems

Networked systems can be large scale, need coordination, information exchange, and data



Much expanded types and severity of uncertainty in networked systems: Data loss, random time delays, network interruption and topology switching, ……

Information and data exchange needs resources and takes time. Information processing and computation will consume computing resources.

# Complexity of Networked Systems

## Data Complexity

Signal Quantization
Data Compression
Big Data Small Information

## Time Complexity

Irregular and Random Sampling
Time-varying and Random Delay
Asynchronous Operation

## Group Complexity

Diversity in Members
Goal Disparity

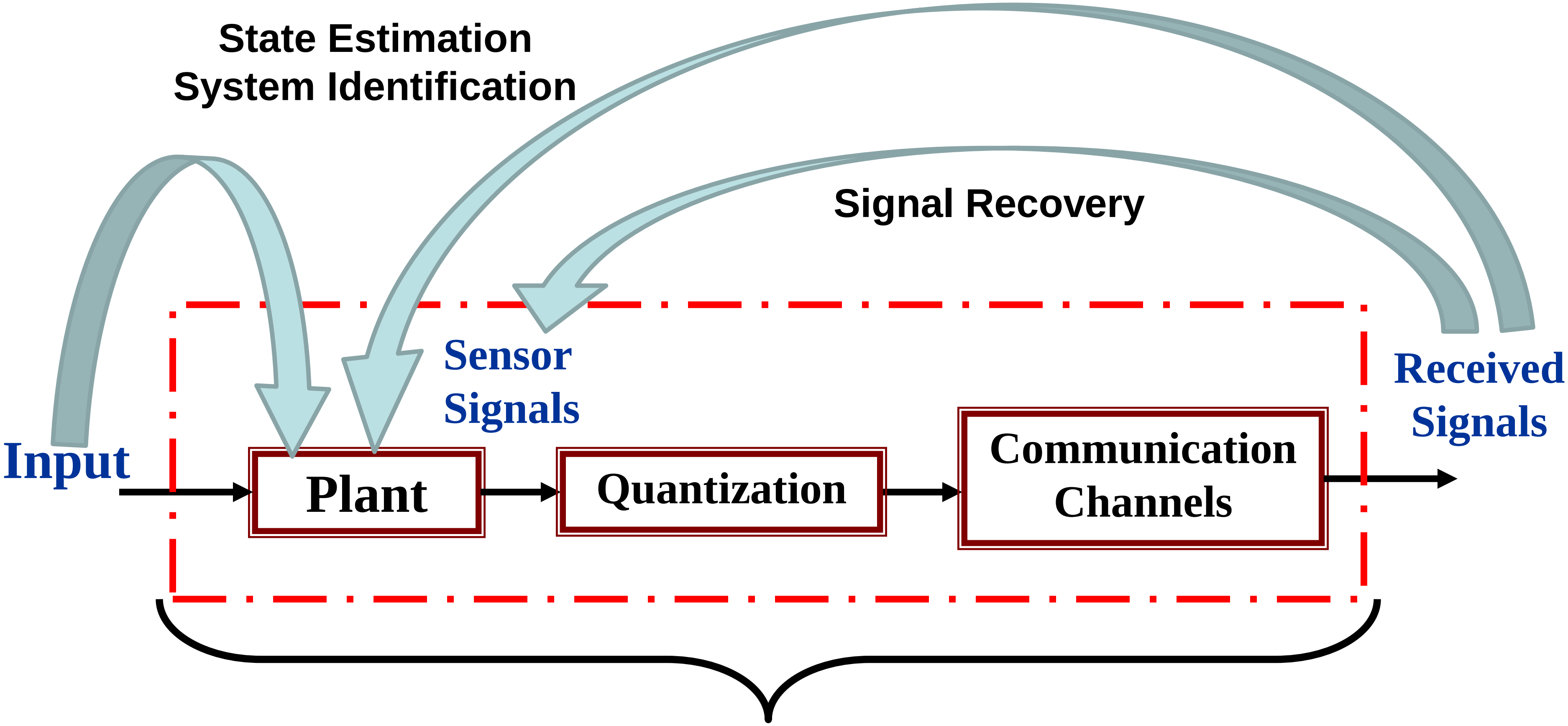
## Spatial Complexity

Data Locations are Distributed
Physically and in Cyber Space
Distributed Computation

# (1) Data Complexity:
# Data Size Reduction

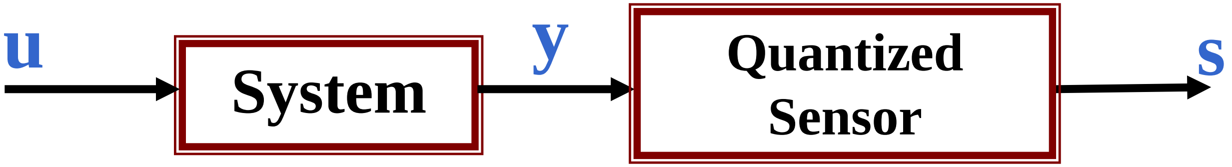
## Estimation, Identification, Learning
## Under Communication Quantization

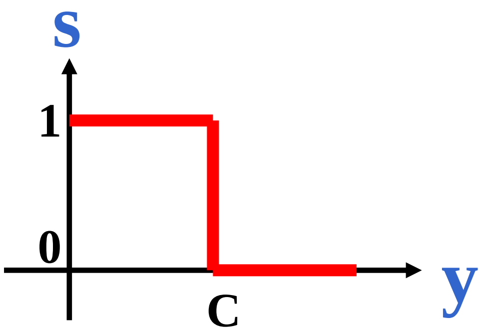
# System Configuration under Quantization



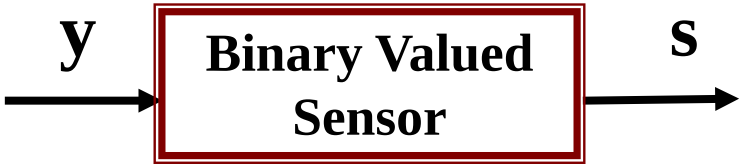
Signal recovering, state estimation, identification
with quantized observations

# Deterministic Framework





## Binary-Valued Sensors

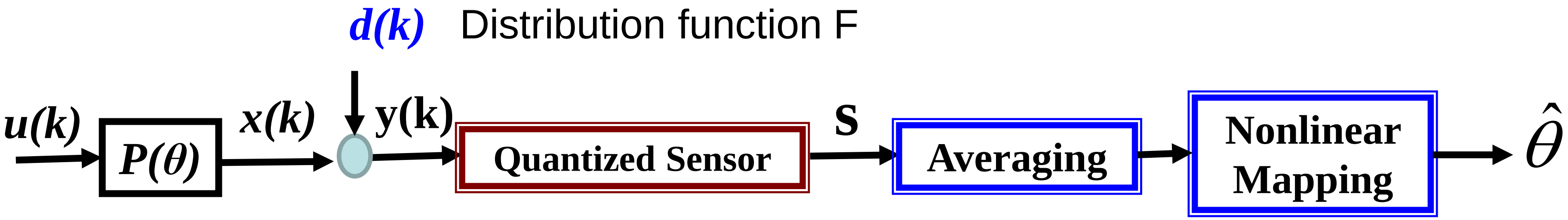


$y > C \longrightarrow S=00000000….$

$y < C \longrightarrow S=11111111….$

**Not Sufficient Information to Know y!**

# Fundamental Solution: Randomize the Observations



Fundamental (simple) Case: $u(k)=1,\ P(\theta)=\theta \Rightarrow x(k)=\theta$

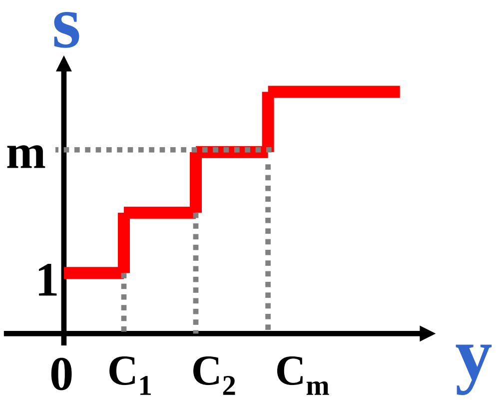
$y_k=\theta+d_k, s_k=S(y_k)$, binary sensor with threshold C

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

**Empirical Measure** $\xi_N=\frac{1}{N}\sum_{k=1}^{N} s_k \ \rightarrow \ p=P\{s_k=1\}=P\{d_k \le C-\theta\}=F(C-\theta)$ **Strong Law of Large Numbers**

Assume invertibility of $F$: $\quad \theta=C-F^{-1}(p)$

$$\hat{\theta}_N=C-F^{-1}(\xi_N)\rightarrow\theta,\ \text{as } N\rightarrow\infty,\ \text{w.p.1}$$



**Quantized Sensors**

$m$ binary sensors $\Rightarrow m$ binary sequences $z_j(k),\ j=1,\dots,m$

$\Rightarrow m$ empirical measures $\xi_j(N)=\frac{1}{N}\sum_{k=1}^{N} z_j(k)$

$\Rightarrow m$ estimates of the same parameter $\hat{\theta}_j(N)=C_j-F^{-1}(\xi_j(N))$

$\Rightarrow m$ strongly (and MS) convergent estimates $\hat{\theta}_j(N)\to\theta,$ w.p.1

$\hat{\theta}(k)=\sum_{j=1}^{m}\alpha_j\hat{\theta}_j(k),\qquad \sum_{j=1}^{m}\alpha_j=1$ Quasi Convex Combination Estimate

Choose $\alpha_j$ Optimally $\Rightarrow$ Optimal QCCE (Markov Estimate)

# Highly Desirable Results

(1) $\hat{\theta}(k) \to \theta, \quad$ w.p.1

and strong convergence rate has been established

(You cannot fail (almost surely)!)

(2) $\hat{\theta}(k) \to \theta,$ in MS and this convergence rate achieves

the Cramer-Rao Lower Bound asymptotically

(Best possible rate (in MS)!) ⟹ **Fundamental for Complexity Analysis**

(3) Asymptotic normality (centered and scaled estimates)

has been established

(Simple for analysis (in distribution)!)

(4) Large (and moderate) deviation principles have been established

(You can characterize reliability accurately (in probability)!)

# Under Periodic Inputs (for achieving the CR Lower Bound and Feedback Invariance)

**Joint Estimation of Systems and Unknown Noise Distributions**

**Extension to Nonlinear Wiener and Hammerstein Models**

**Error Probability with Large Deviation Principles**

**System Complexity Analysis**

**Extension to General Nonlinear Systems**

- L.Y. Wang, J.F. Zhang, G. Yin, System Identification Using Binary Sensors, ***IEEE Trans. Automat. Contr.,*** Vol. 48, pp. 1892-1907, 2003.
- Le Yi Wang, George Yin, Ji-feng Zhang, Yanlong Zhao, System Identification with Quantized Observations, Boston, MA: Birkhuser, 2010

Systems & Control: Foundations & Applications

Le Yi Wang
G. George Yin
Ji-Feng Zhang
Yanlong Zhao

System Identification
with Quantized
Observations

## Extension to General Inputs (Still Achieving the CR Lower Bound Asymptotically)

- Jin Guo, Le Yi Wang, George Yin, Yanlong Zhao, Jifeng Zhang, Asymptotically efficient identification of **FIR systems** with quantized observations and **general quantized inputs**, ***Automatica***, Vol. 57, pp. 113-122, 2015.
- Jin Guo, Le Yi Wang, George Yin, Yanlong Zhao, Ji-Feng Zhang, Identification of **Wiener Systems** with **Quantized Inputs** and Binary-Valued Output Observations, ***Automatica***, Vol. 78, pp. 280-286, 2017.

## Recent Development: Yanlong Zhao' Group: Data fusion using different sensors (Still Achieving the CR Lower Bound Asymptotically)

# (2) Time Complexity
# Data Frequency Reduction

**Transmit More Information in Data**

# Sampling and Quantization Combination

Input

System

System Output

Sampling & Quantization

Communication Channels

Received Signals

## Question on Information vs Complexity

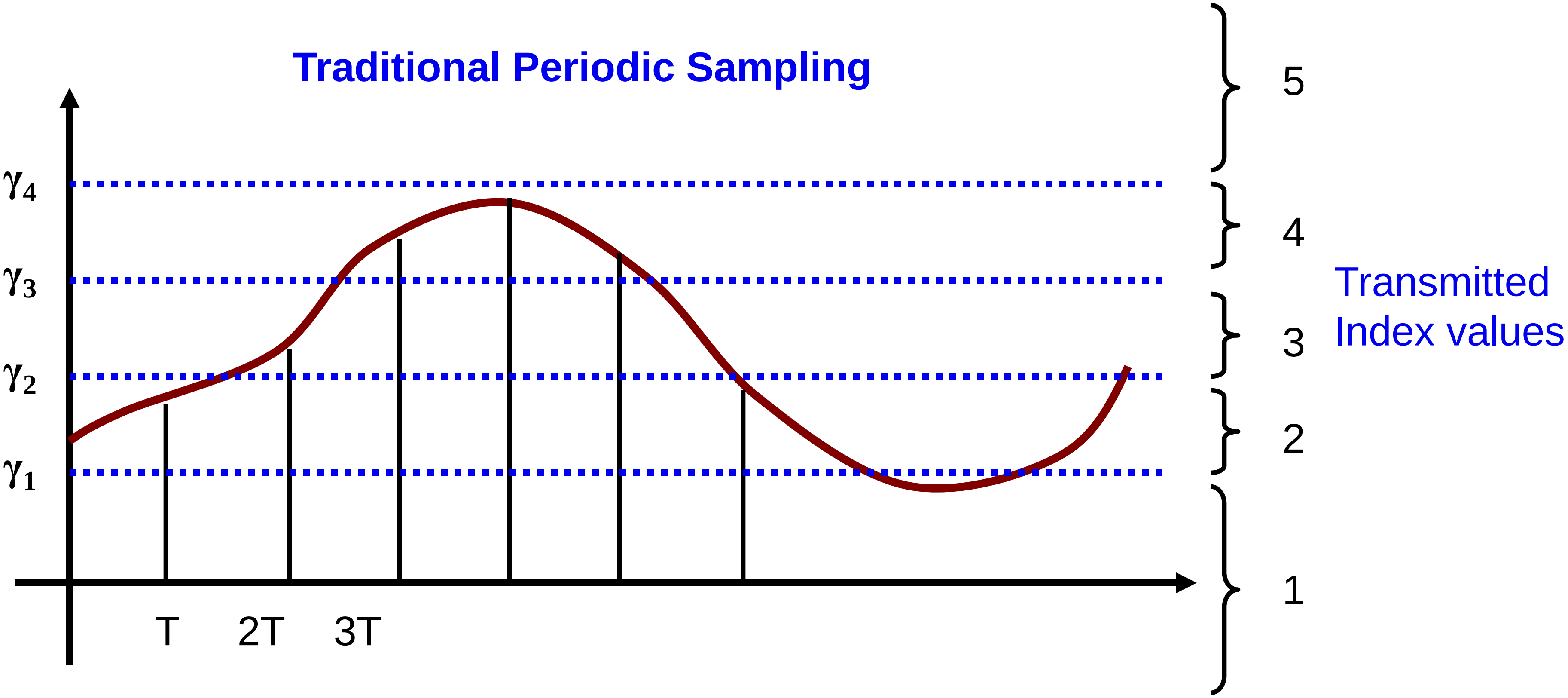
How much information about a signal that can be obtained after sampling and quantization?

**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.

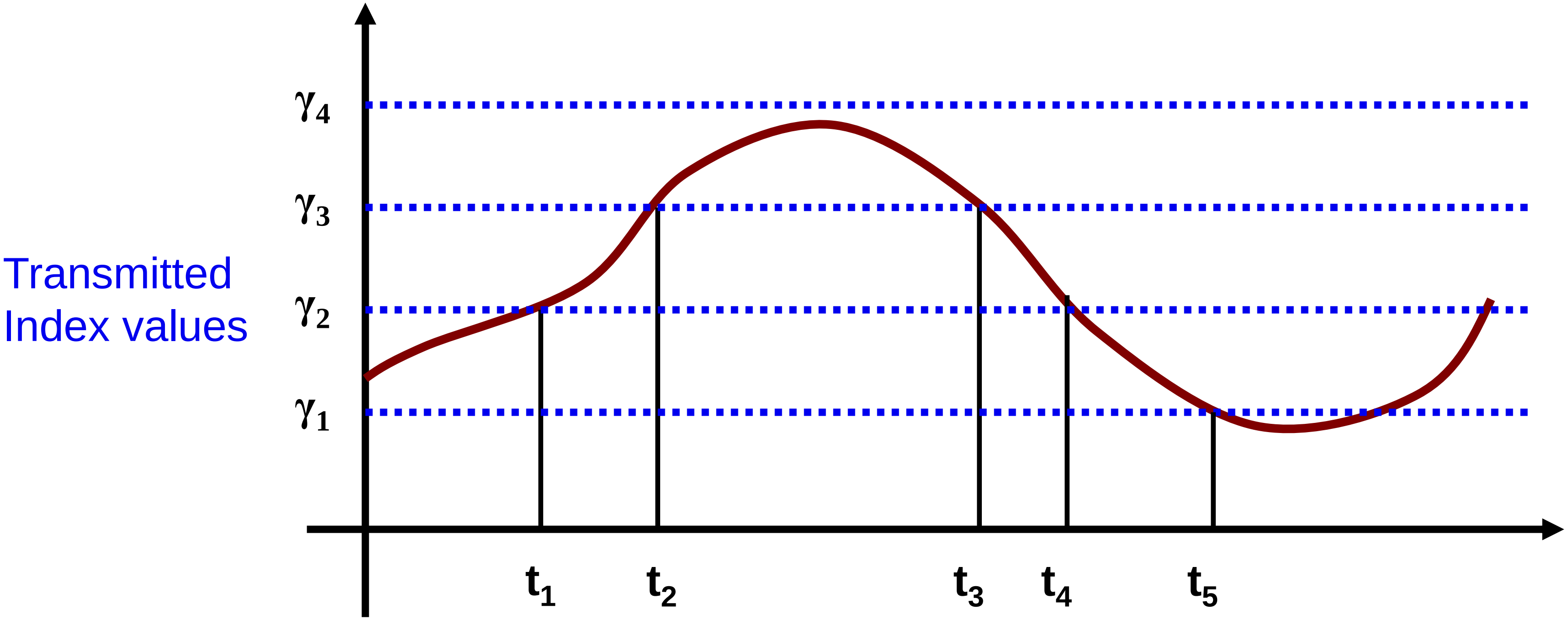
**Traditional Periodic Sampling**



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.**

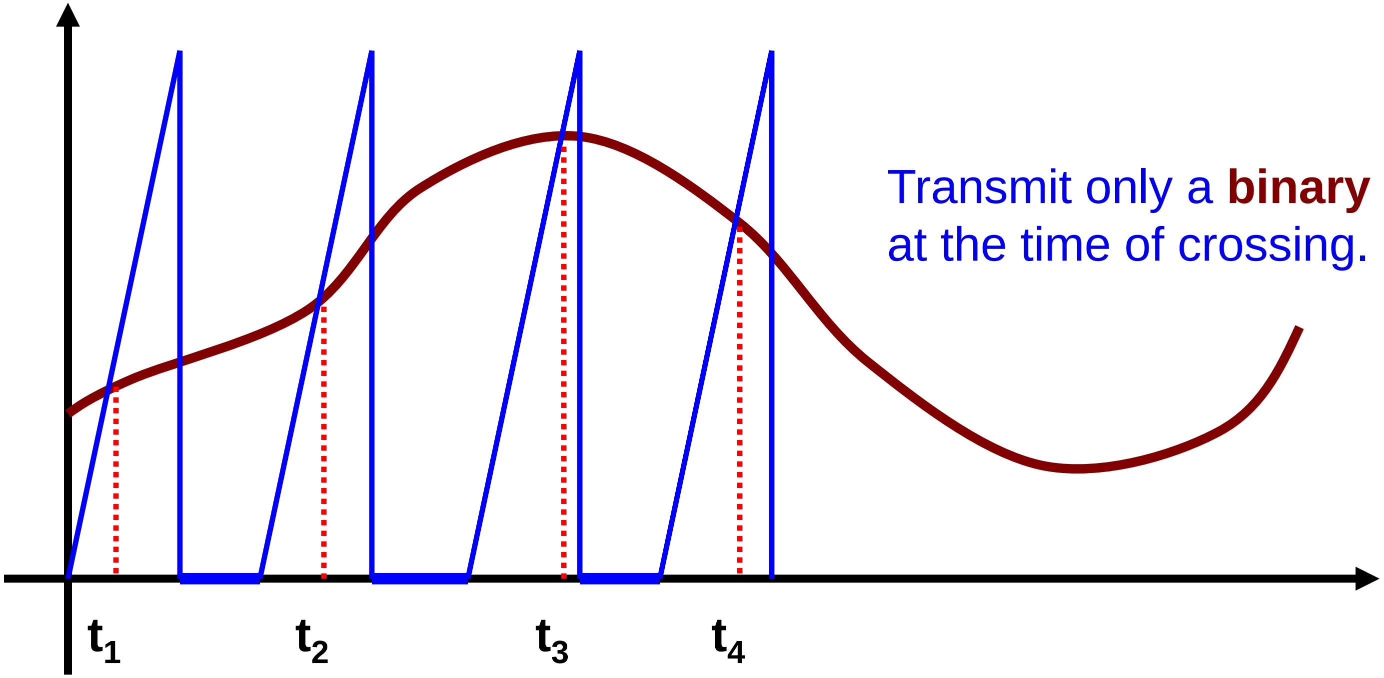
# Event Triggered Sampling



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.**

# PWM-Based Sampling



Transmit only a **binary indicator** at the time of crossing.

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 or delays.

**Effective sampling scheme, control of sampling density, issues of irregular sampling with state estimation and control need to be resolved.**

# Fundamental Question:
# Will irregularly sampled data provide sufficient information?

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned}$$

## Observability

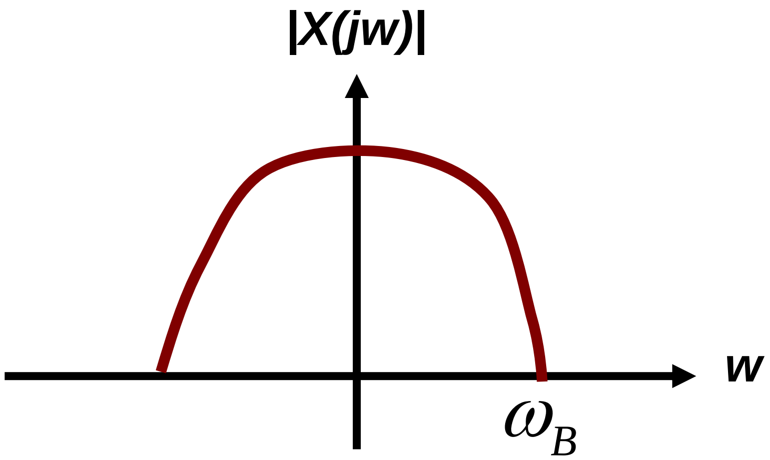
$d_k = 0.\ \ N \geq n$

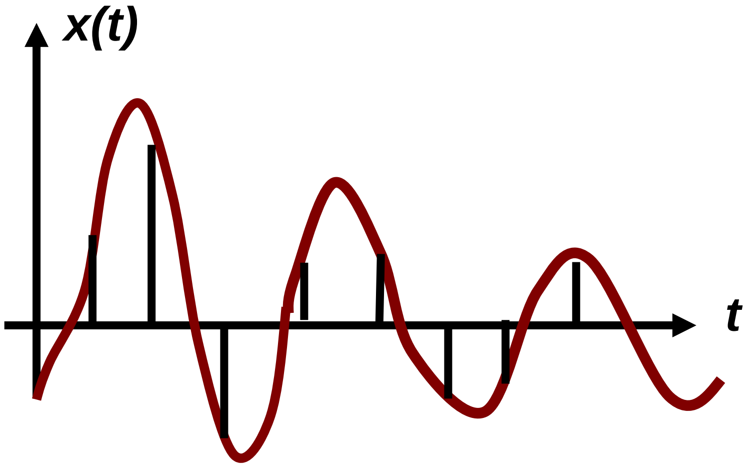
Will $t_1, t_2, \ldots, t_N,$ and $z(t_k) = y(t_k)$ be sufficient for estimating $x(0)$?

**In general, the answer is No!**

**Can we use Shannon’s Sampling Theorem to analyze this?**

# Shannon-Nyquist's Sampling Theorem for Signals





1. 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
(Fundamental Complexity Relationship)

# Why Shannon's Sampling Theorem cannot be applied?

1. Under any initial condition, if the corresponding *y(t)* is not zero, it always has **unbounded bandwidth**.
2. Sampling time is **not periodic**.
3. Shannon's Sampling Theorem requires an "infinite" data set for exact reconstruction. We only have a **finite number** of sampling points.
4. Shannon's Sampling Theorem is **not causal**: You must collect all data first.
5. Observability (and controllability, identifiability, etc.) is an **exact statement on the system**, no approximation is allowed.

**We need a new sampling theorem for systems!**

# Our New General Sampling Theorem for Systems

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

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

tight, cannot be improved

## Theorem

Suppose the system is observable,

$0 \le t_i \le T, i = 1, \ldots, N$ and $N > \mu_T$. Then $\Phi_{\mathrm{N}}$ is full rank.

Asymptotically, $\dfrac{\mu_T}{T} \approx \dfrac{\omega}{\pi} = \mu,$

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.

Le Yi Wang, Chanying Li, George Yin, Lei Guo, Chengzhong Xu, State Observability and Observers of Linear-Time-Invariant Systems under Irregular-Sampling and Sensor Limitations, ***IEEE Transactions on Automatic Control***, 56, no. 11, pp. 2639 - 2654, 2011.

# This Sampling Theorem for LTI Systems

1. **For linear time invariant systems**
2. **Irregular sampling**
3. **Finite Data**
4. **Causal state reconstruction**
5. **Critical complexity relationship**

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

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

**Sampling density**

**system characteristic bandwidth**

**Fundamental Complexity Relationship for Systems**

# This complexity relationship is essential for

## Extension to Convergence, Convergence Rates, Estimation Accuracy

Le Yi Wang, Chanying Li, George Yin, Lei Guo, Chengzhong Xu, State Observability and Observers of Linear-Time-Invariant Systems under Irregular-Sampling and Sensor Limitations, ***IEEE Transactions on Automatic Control***, 56, no. 11, pp. 2639 - 2654, 2011.

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

Le Yi Wang, Wei Feng, George Yin, Joint State and Event Observers for Linear Switching Systems under Irregular Sampling, ***Automatica***, 49, pp. 894-905, 2013.

## Extension to System Identification

Biqiang Mu, Jin Guo, Le Yi Wang, George Yin, Lijian Xu, Wei Xing Zheng, Identification of linear continuous-time systems under irregular and random sampling, ***Automatica***, Vol. 60, pp. 100-114, 2015.

## Extension to Controllability

Ping Zhao, Le Yi Wang, George Yin, Controllability and adaptation of linear time-invariant systems under irregular and Markovian sampling, ***Automatica***, Vol. 63, pp. 92-100, January 2016.

# (3) Group Complexity
# Data Total Volume Reduction

## Decision and Complexity Based System Identification

# Traditional System Identification

Example: $$y_k = \phi_k^T \theta + d_k, \qquad k = 1, \ldots, N$$

**Algorithms**: $$\hat{\theta}_N = \left( \Phi_N^T \Phi_N \right)^{-1} \Phi_N^T Y_N, \; \Phi_N = \begin{bmatrix} \phi_1^T \\ \vdots \\ \phi_N^T \end{bmatrix}, \quad Y_N = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$$

**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.**

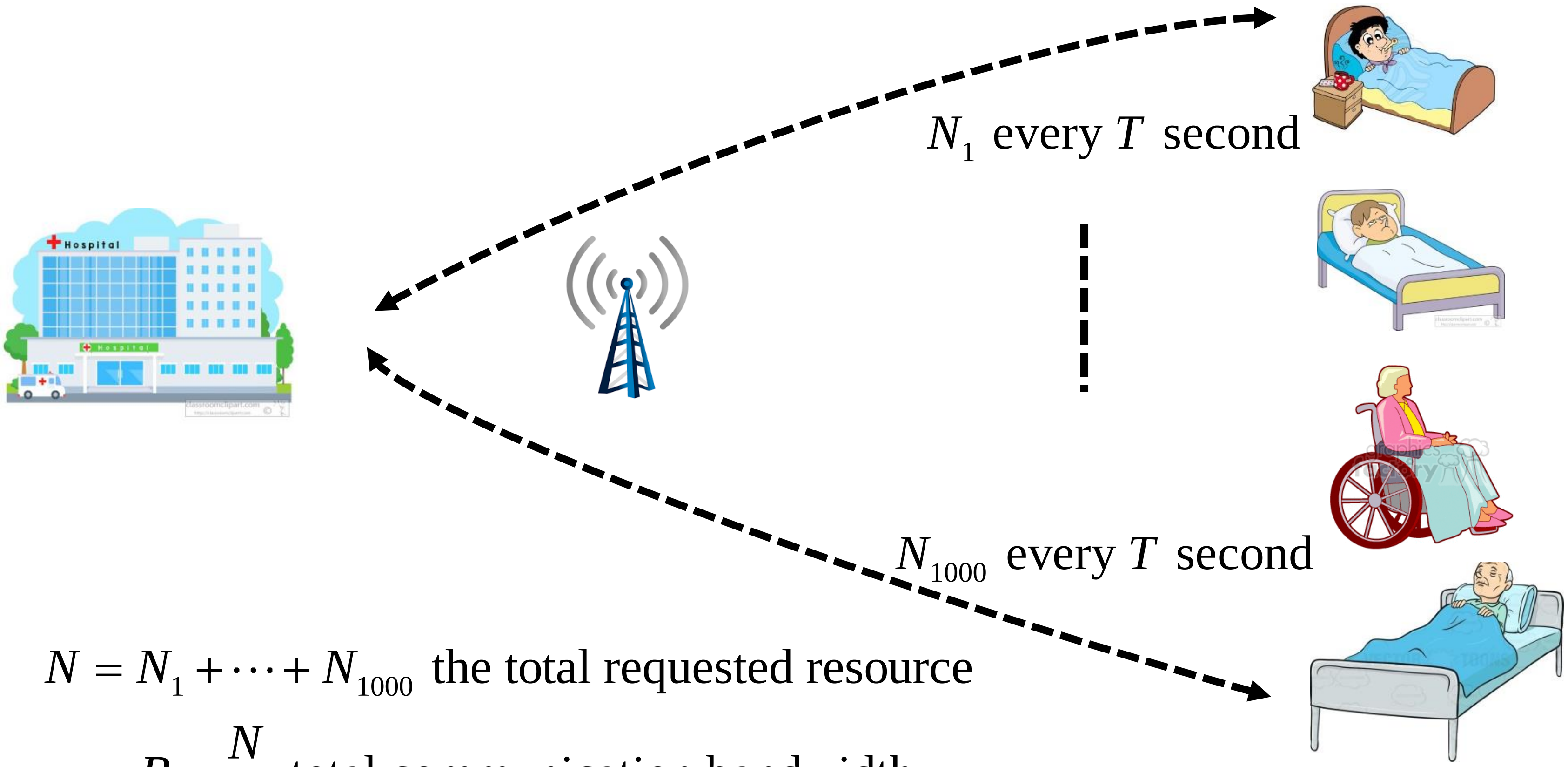
## Example 1: Robust Feedback Controller

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.

# Telemedicine: Connected Patients and Remote Automated Group Monitoring



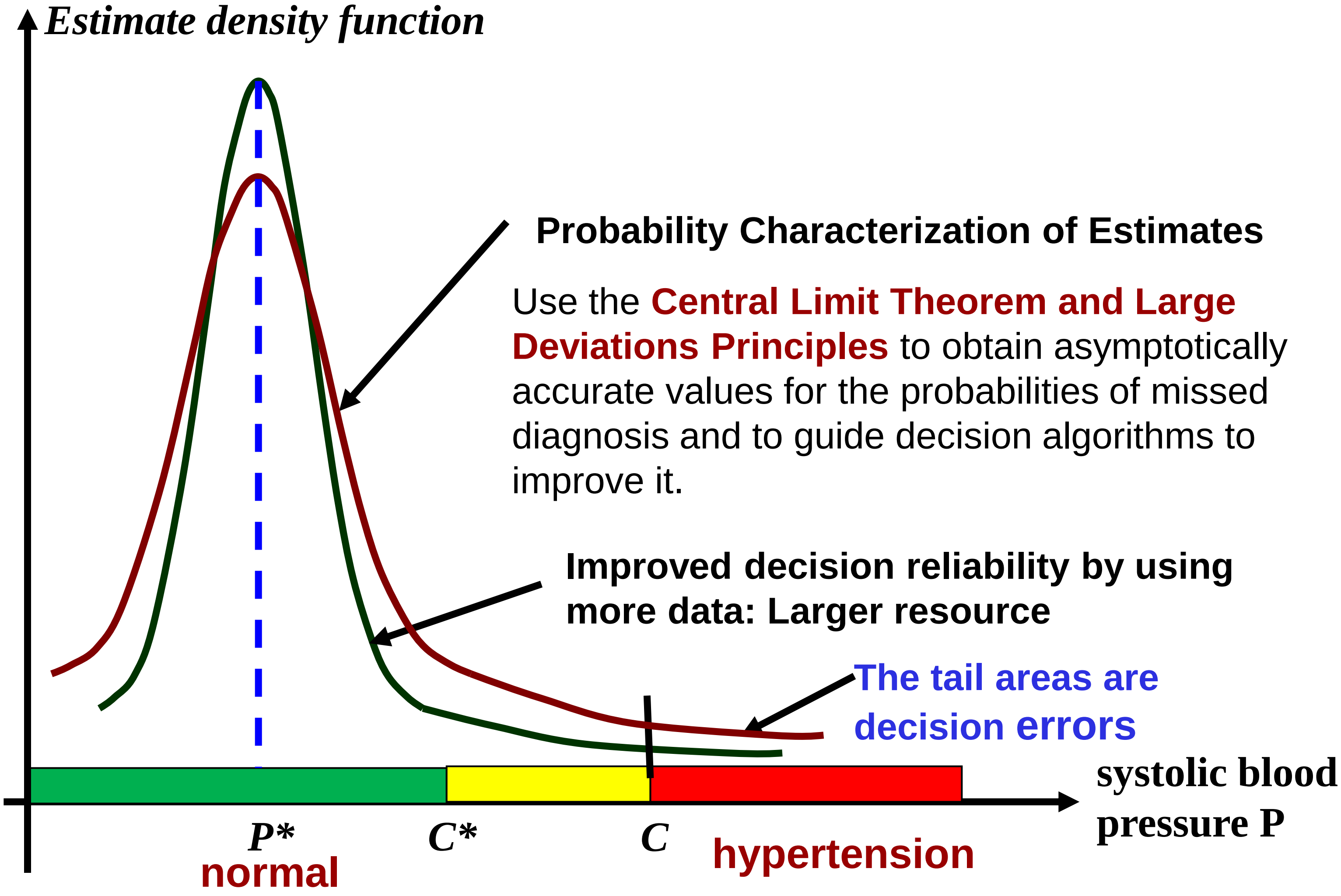
$$N = N_1 + \cdots + N_{1000} \text{ the total requested resource}$$

$$B = \frac{N}{T} \text{ total communication bandwidth}$$

## Resource Allocation Strategy 1

**All users are assigned the equal number of data points in communication during the updating time interval *T*.**

## Is this a good strategy?



**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.

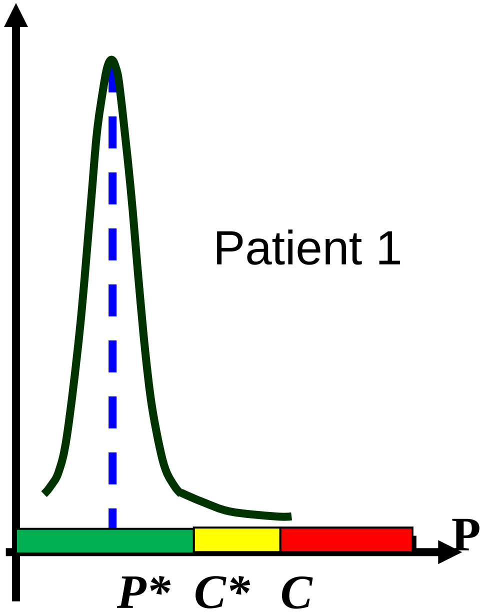
**Improved decision reliability by using more data: Larger resource**

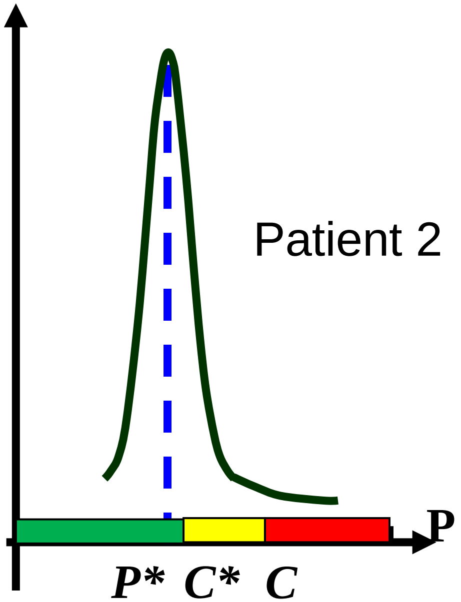
**The tail areas are decision errors**

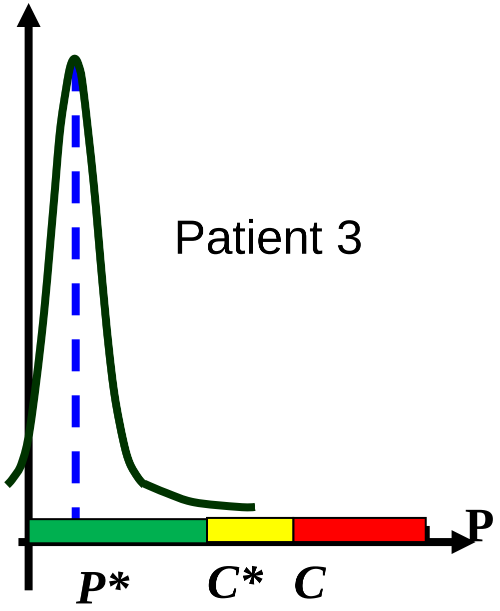
**Blood Pressure Monitoring Example**

Decision Error Probability (for a normal patient)

$$\alpha(p_{true}, N) = P\left\{\hat{p}_N - p_{true} > C \middle| p_{true}\right\} \Leftarrow \text{ Large Deviation Principle}$$







**Equal Resource Allocation (Population Based)** → **Large Reliability Variations (Individual Reliability)**

- The overall reliability is determined by the worst case
- Many resources are wasted

**Uniform Reliability** → **Dynamic Resource Allocation (Individualized and Need Algorithms)**

- The uniform reliability
- Resources are saved

$$\text{For } \alpha(p_{true}, N) = \alpha_0 \Rightarrow N = f(p_{true})$$

The true parameter is unknown, so we need to estimate N

# 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, $k \to \infty$

**All these properties have been established for individual parameters under Gaussian i.i.d. cases**

1.L.Y. Wang, G. Yin, J. Guo, B.Q. Mu, L.J. Xu, From Wiener filtering to recent advances on complexity based system identification and state estimation, **IEEE Conference on Norbert Wiener for the 21st Century**, Boston, June 24-26, 2014.

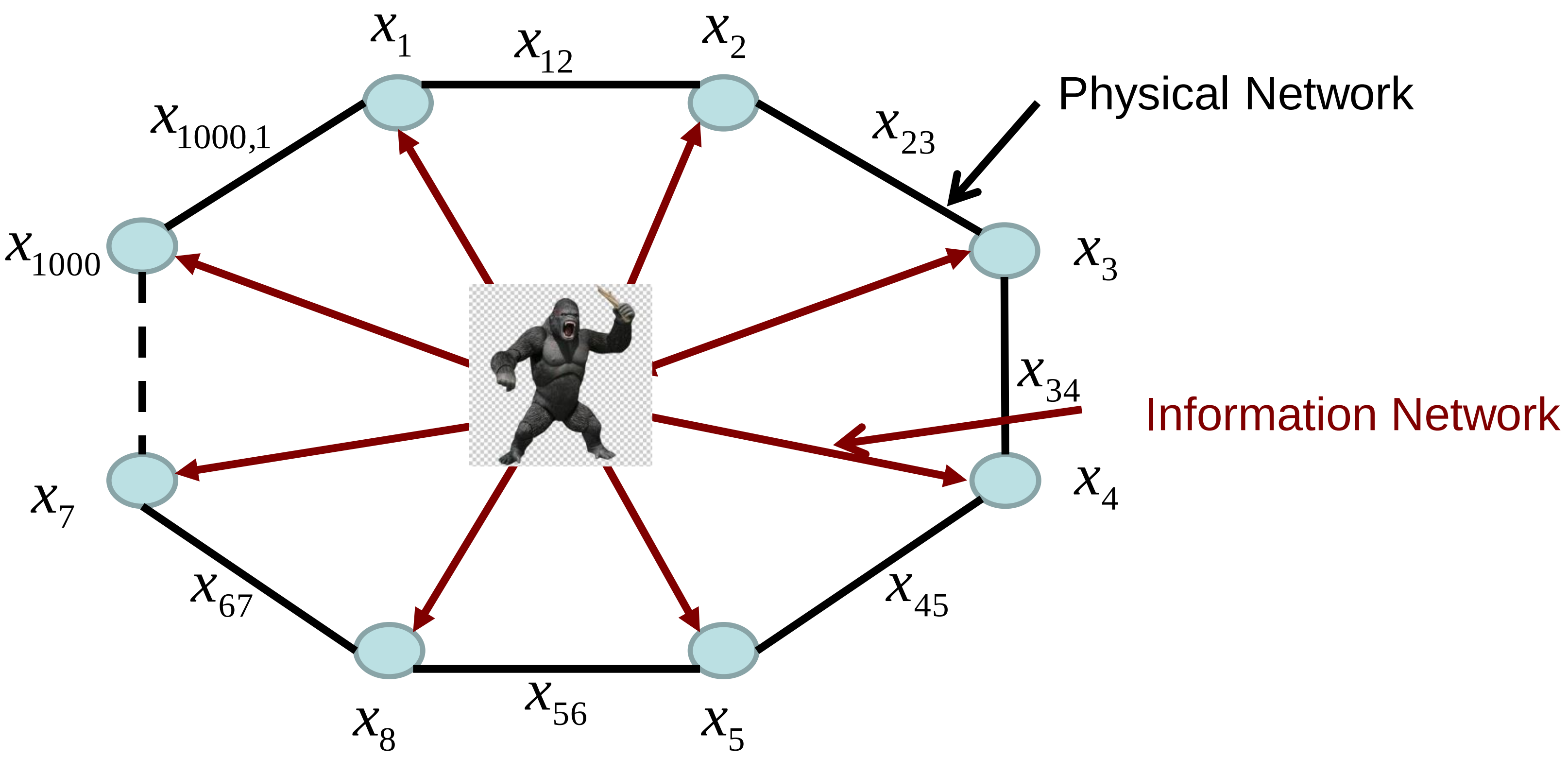
2.Jin Guo, Bi-Qiang Mu, Le Yi Wang, George Yin, Lijian Xu, Decision-based system identification and adaptive resource allocation, ***IEEE Transactions on Automatic Control***, 62-5, pp. 2166-2179, 2017**.**

Le Yi Wang, Yanlong Zhao, Ting Wang, Robust control and system identification: A complexity perspective. Sci Sin Math, 2016

# (4) Spatial Complexity
# Data Distance Reduction

## Distributed Optimization
## In Networked Systems

# Centralized Strategy: One Node is the King



The size of the network: $x = [x_1, x_{12}, x_2, x_{23}, x_3, \ldots, x_{1000}]$

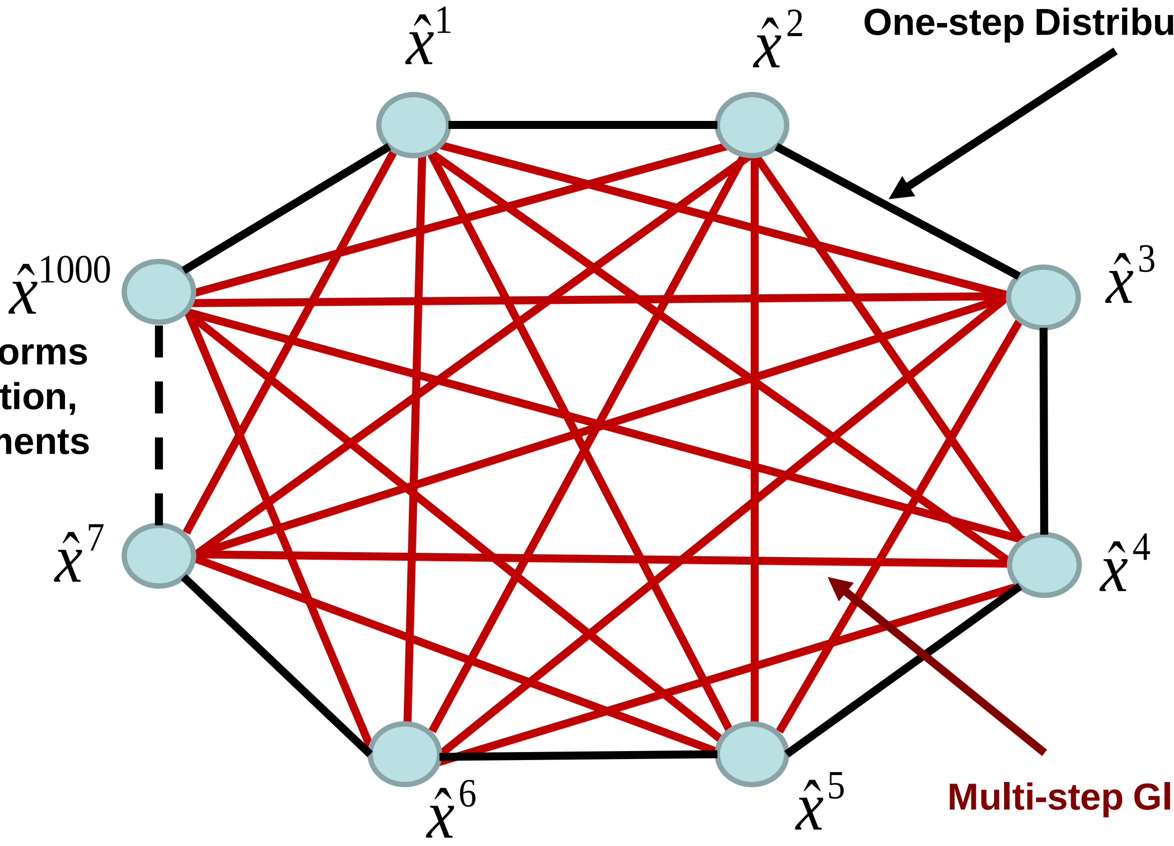
- **The King communicates and computes the whole (optimal) state (high complexity in one node)**
- **The King knows everything (node privacy is lost to the King)**

**It is actually a simple structure, not very high complexity.**

**But it is vulnerable!**

# Non-Strict Distributed Strategy: Every Node is a King

The size of the network: $x = [x_1, x_{12}, x_2, x_{23}, x_3, \ldots, x_{1000}]$



**One-step Distributed Consensus**

**Each node performs global optimization, but only implements one variable.**

**Multi-step Global Information**

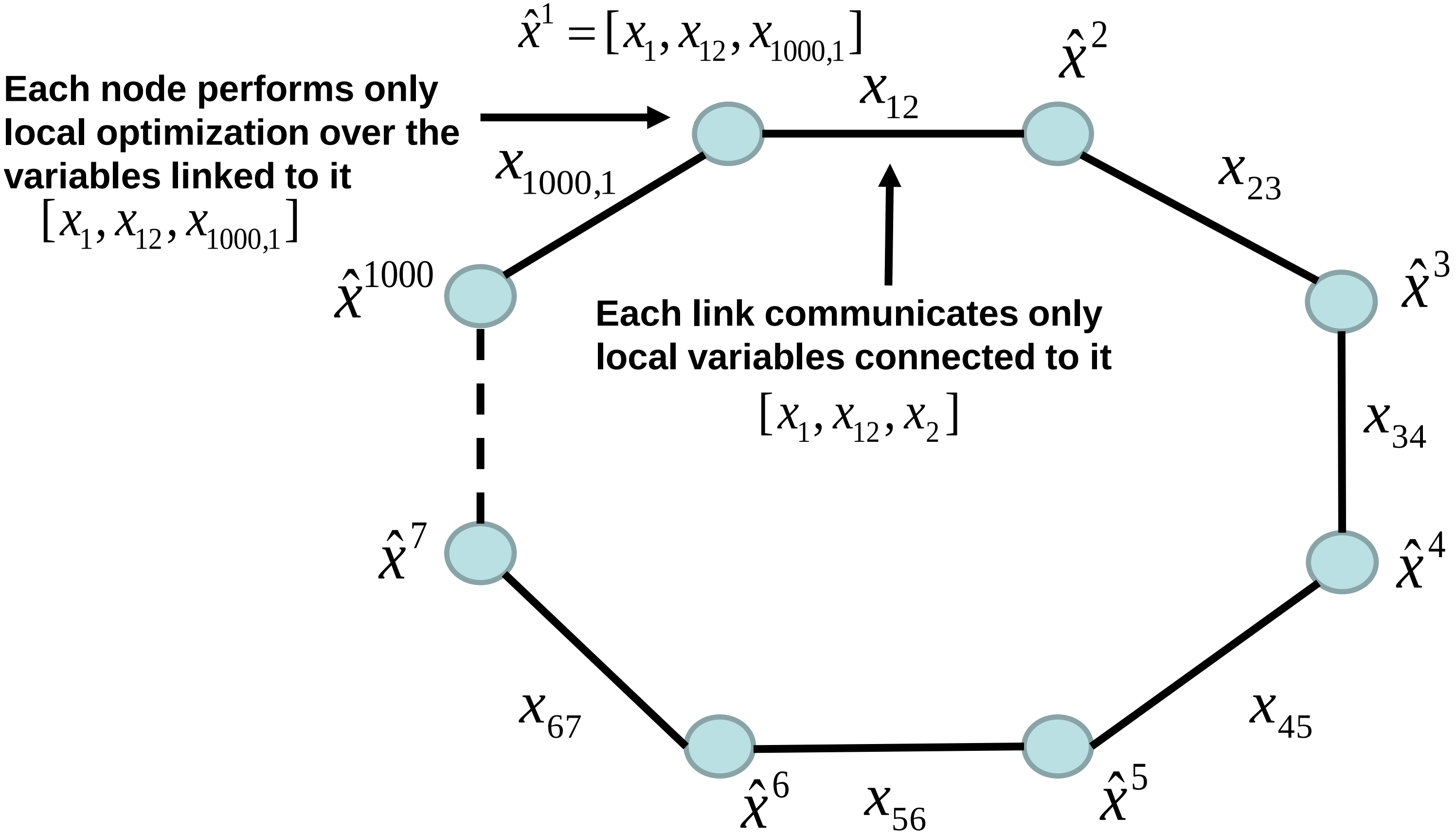
**Now, every subsystem has the global information!**

**No loss of capability:** **What can be achieved by a central strategy can always be achieved by this “multi-hop” distributed strategy**

**It is actually a complicated structure, very high complexity.**

**Much More Expensive than the Centralized Strategy!**

# Strictly Distributed Approach: No One is a King



**Each node performs only local optimization over the variables linked to it**
$[x_1, x_{12}, x_{1000,1}]$

**Each link communicates only local variables connected to it**
$[x_1, x_{12}, x_2]$

**Low communication and computing complexity, high resiliency, good privacy**

**Imposing constraints on network structures and achievable performance, but many network systems can achieve that!**

# Conditions of Strictly Distributed Approach

**(1) Privacy**: Subsystems use and maintain only their own and neighbors information. No multi-hop information passing is allowed.

**(2) Local Optimization**: Subsystems perform only local calculations (e.g., local gradient/subgradient) over variables linked to them, independent of the network size.

**(3) Local Communication**: Communication channels carry only data that are linked to it. The data size on each channel is independent of the network size.

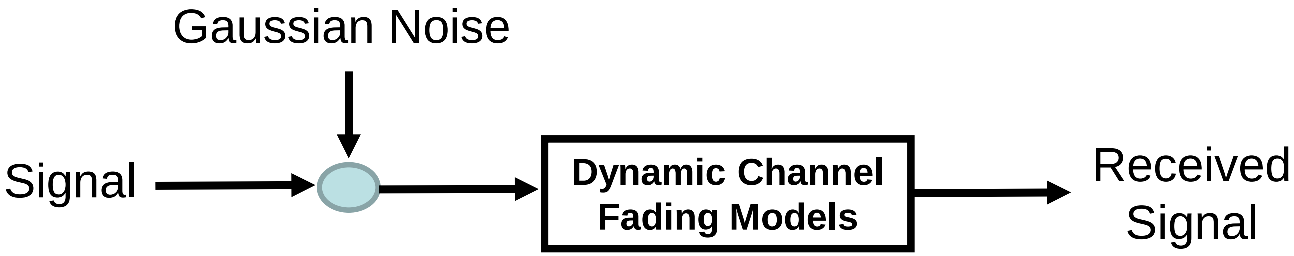
**(4) Scalability**: Adding/deleting a non-neighbor player to the network should not affect its control and decision (it is not aware of this addition).

- Le Yi Wang, Shu Liang, Siyu Xie, George Yin, Masoud Nazari, **Strictly distributed** optimization for cyber-physical networks, **in preparation**
- Shu Liang, Le Yi Wang, George Yin, Exponential convergence of a **distributed** primal-dual convex optimization algorithm without strong convexity, ***Automatica***, accepted in 2019.
- Shu Liang, Le Yi Wang, and George Yin, **Distributed** Dual Averaging with Iterate-Averaging Feedback for Nonsmooth Convex Optimization, ***Automatica***, pp.101:175-181, 2019.
- Shu Liang, Le Yi Wang, George Yin, **Distributed** Dual Subgradient Algorithms with Iterate-Averaging Feedback for Convex Optimization with Coupled Constraints, ***IEEE Transactions on Cybernetics***, accepted in 2019.

# Communication Uncertainty and a Stochastic Approximation Framework to Deal with Control-Communication Co-Design

# Communication System Reality

## Desirable Channel Models:



Gaussian Noise

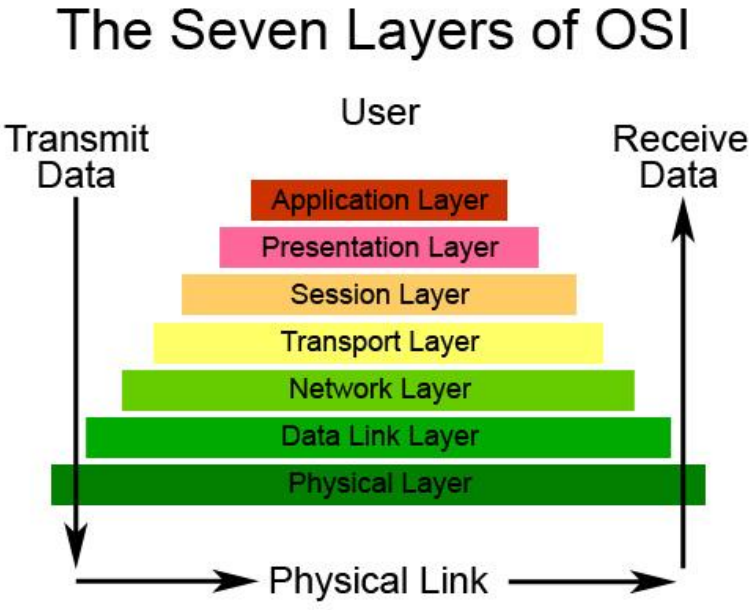
Signal

**Dynamic Channel Fading Models**

Received Signal

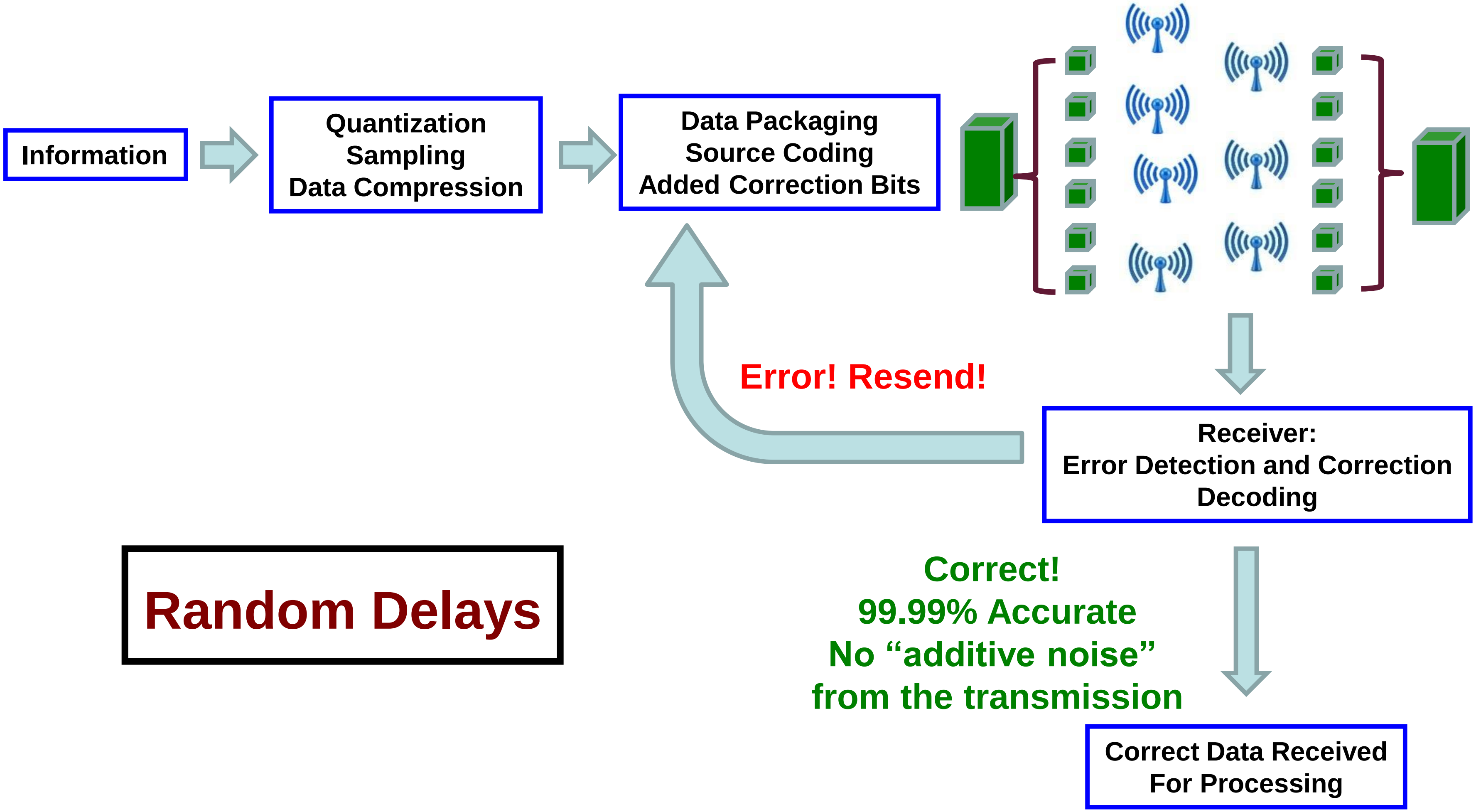
**Very reasonable and commonly used physical layer and analog channel model**

## Open Systems Interconnection Model (OSI Model)



**Control/Communication Interfacing Application Layer: Digital Communication Network**

# Scenarios from Communication Systems



Information

Quantization
Sampling
Data Compression

Data Packaging
Source Coding
Added Correction Bits

Error! Resend!

Receiver:
Error Detection and Correction
Decoding

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

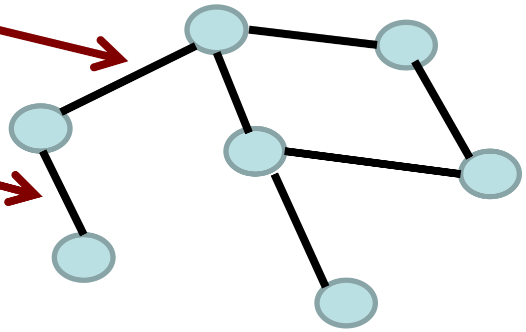
Correct Data Received
For Processing

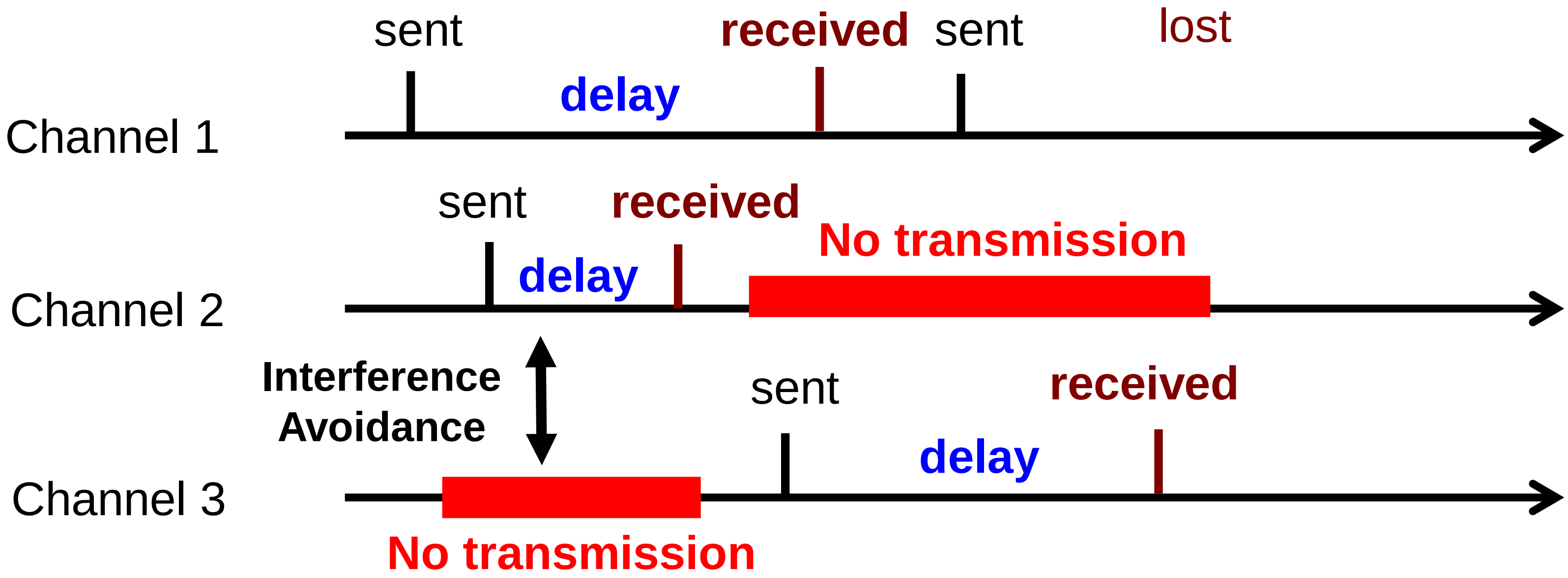
**Random Delays**

# Communication Network Reality

- **Channels can interrupt randomly**
- **To avoid signal interference, Channels must operate in a certain order**
- **Signals arrive at different time**

1. **Randomly Switching Topology**
2. **Random Delays**
3. **Random Sampling and Asynchronous Operation**





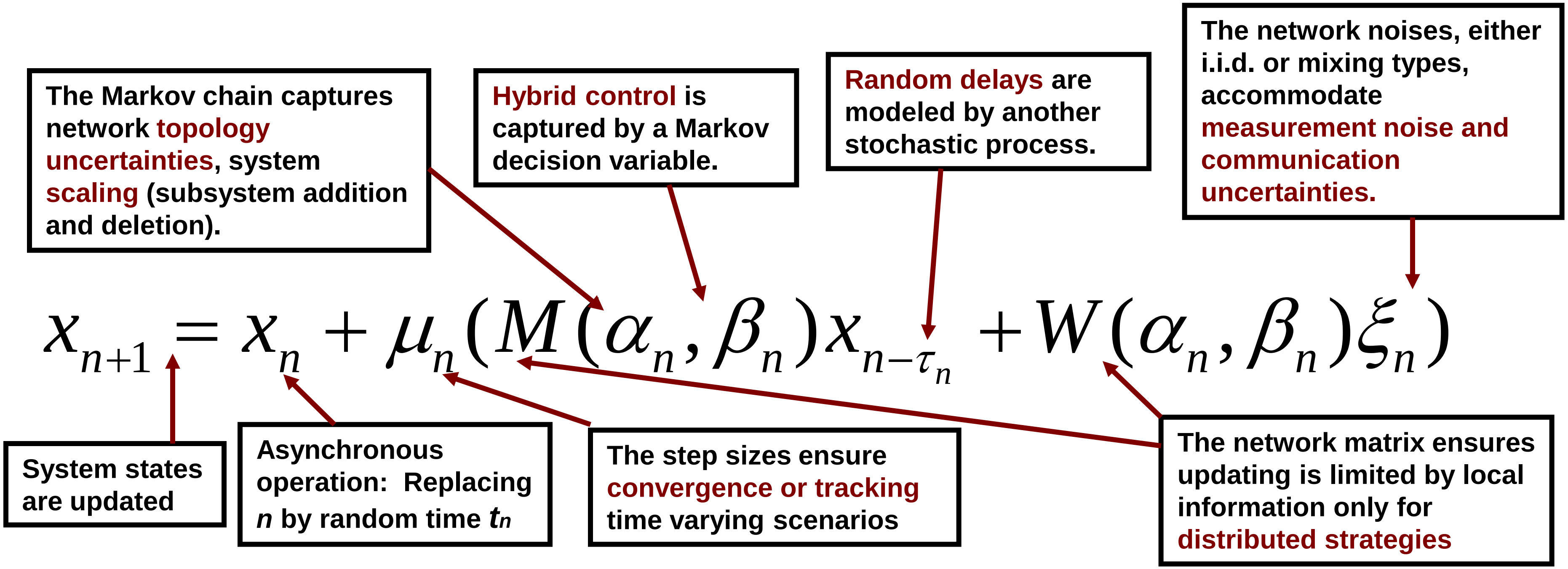
Channel 1: sent, **delay**, **received**, sent, lost

Channel 2: sent, **delay**, **received**, **No transmission**

**Interference Avoidance**

Channel 3: **No transmission**, sent, **delay**, **received**

# How can an SA Structure Accommodate Unique Network System Features?



The Markov chain captures network **topology uncertainties**, system **scaling** (subsystem addition and deletion).

**Hybrid control** is captured by a Markov decision variable.

**Random delays** are modeled by another stochastic process.

The network noises, either i.i.d. or mixing types, accommodate **measurement noise and communication uncertainties.**

$$x_{n+1} = x_n + \mu_n (M(\alpha_n, \beta_n) x_{n-\tau_n} + W(\alpha_n, \beta_n) \xi_n)$$

System states are updated

Asynchronous operation: Replacing *n* by random time $t_n$

The step sizes ensure **convergence or tracking** time varying scenarios

The network matrix ensures updating is limited by local information only for **distributed strategies**

## Some Appealing Features:

1. **Extensive Convergence Results**:

   Strong convergence, MS convergence, asymptotic normality, and the related convergence rates

2. Suitable **Complexity Analysis** (best possible rate of convergence)

   With Post-Iterate Averaging and under some conditions, the MS convergence is asymptotically efficient (reaching the **Cramer-Rao Lower Bound**)

3. **Regime Switching** results are available.

4. The limiting ODE or SDE can be used to analyze stability and performance under **random delays**.

H.J. Kushner and G. Yin, Stochastic Approximation and Recursive Algorithms and Applications, 2nd Edition, Springer-Verlag, New York, 2003, [Applications of Mathematics, Volume 35].

# Fundamental Mathematics Problem

$$x_{n+1} = x_n + \mu_n(M(\alpha_n)x_{n-\tau} + W(\alpha_n)\xi_n)$$

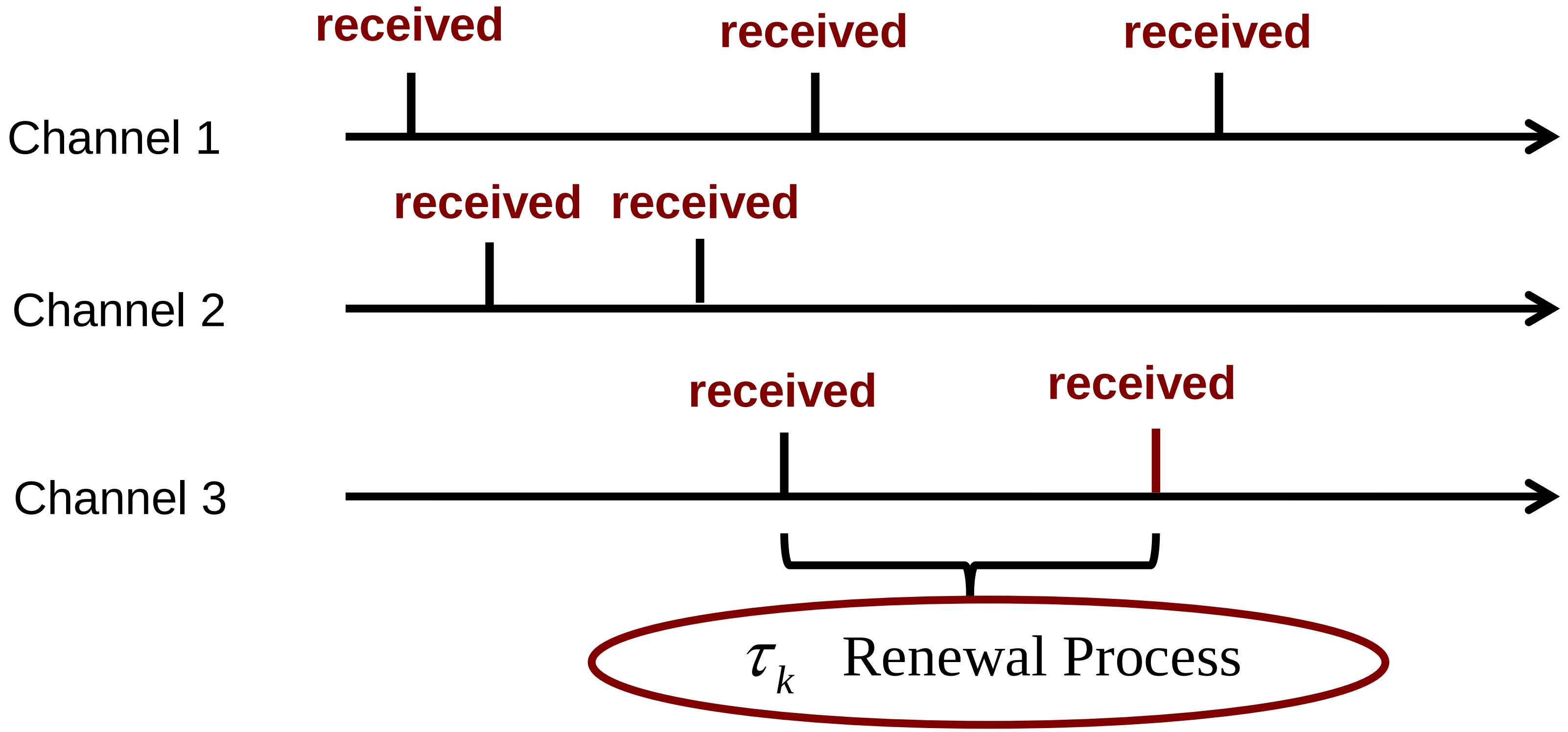
**Limit Stochastic Functional Differential Equations:**
**Three stochastic processes: Brownian Motion, Markov Chain, Random Delays**

$$dX(t) = f(\alpha(t), X_t)dt + g(\alpha(t), X_t)dw(t)$$

$$X_t(\lambda) = \{X(t+\lambda) : -\tau \le \lambda \le 0\}$$

**Very difficult problems to analyze and solve, especially before 2009.**

# Asynchronous Operation



Channel 1

received received received

Channel 2

received received

Channel 3

received received

$\tau_k$ Renewal Process

- G. Yin, Q. Yuan, L.Y. Wang, **Asynchronous** stochastic approximation algorithms for networked systems: **Regime-switching** topologies and multi-scale structure, ***Multiscale Modeling and Simulation***, Vol. 11, No. 3, pp. 813–839, 2013.

# Switching Network Topology

**Consider the simple case of no-delay and linear (fast) switching systems:**

$$dX(t) = A(\alpha(t))X(t)dt + \sum_{j=1}^{d} B_j(\alpha(t))X(t)dw^j(t) \qquad (1)$$

$\alpha(t) \in M = \{1, \ldots, m\}$ is a continuous-time, finite state Markov chain.

If $\alpha(t)$ is positive recurrent, then there exists a stationary density

$$v_i > 0, i \in M, \sum_{i=1}^{m} v_i = 1.$$

The Brownian motion $w^j(t),\ j = 1, \ldots, d$ are mutually independent.

$\lambda^{A}_{\max}(i)$ (and $\lambda^{A}_{\min}(i)$) $=$ the largest (and smallest) eigenvalue of $\frac{1}{2}(A(i) + A'(i))$

$\lambda^{B_j}_{\max}(i)$ (and $\lambda^{B}_{\min}(i)$) $=$ the largest (and smallest) eigenvalue of $\frac{1}{2}(B_j(i) + B_j{}'(i))$

$\eta^{j}_{\max}(i)$ (and $\eta^{j}_{\min}(i)$) $=$ the largest (and smallest) eigenvalue of $B_j(i)B_j{}'(i)$

- R.Z. Khasminskii, C. Zhu, **G. Yin**, Stability of regime-switching diffusions, ***Stochastic Processes and their Applications***, pp. 1037-1051, August 2007.
- **G. Yin** and C. Zhu, Hybrid Switching Diffusions: Properties and Applications, Springer, New York, 2010, [Stochastic Modeling and Applied Probability, Volume 63]

**Weighted Average Markovian Switching**

**From Drift Part similar to deterministic systems**

**From Diffusion Part unique to stochastic systems Ito Formula**

$$\mu_{\max} = \sum_{i=1}^{m} v_i [\lambda_{\max}^{A}(i) + \sum_{j=1}^{d} \overbrace{(\eta_{\max}^{j}(i) - 2(\lambda_{\min}^{B_j}(i))^2)]}]$$

$$\mu_{\min} = \sum_{i=1}^{m} v_i [\lambda_{\min}^{A}(i) + \sum_{j=1}^{d} (\eta_{\min}^{j}(i) - 2(\lambda_{\max}^{B_j}(i))^2)]$$

# Theorem

(a) If $\mu_{\max} < 0,$ then the siwtching diffusion (1) is asymptotically stable in probability.

(b) If $\mu_{\min} > 0,$ then the siwtching diffusion (1) is unstable.

- **Switching among stable systems may lead to unstable systems**
- **Switching among unstable systems may lead to stable systems**
- **Noise can sometimes assist stability if they are small**

- S.L. Nguyen, D.T. Nguyen, G. Yin, and L.Y. Wang, Modeling and controls of large-scale **switching diffusion networks** with mean-field interactions, 2019 8th International Conference on Systems and Control.
- T. Bui, X. Cheng, Z. Jin, and G. Yin, Approximation of a class of non-zero-sum investment and reinsurance games for **regime-switching jump-diffusion** models, ***Nonlinear Analysis: Hybrid Systems***, 32, 276—293, 2019.
- George Yin, Le Yi Wang, Thu Nguyen, **Switching** Stochastic Approximation and Applications to Networked Systems, ***IEEE Transactions on Automatic Control***, accepted in 2018.
- Thu Nguyen, Le Yi Wang, George Yin, Hongwei Zhang, Shengbo Eben Li, Keqiang Li, Impact of Communication **Erasure Channels** on Control Performance of Connected and Automated Vehicles, ***IEEE Transactions on Vehicular Technology***, Vol. 67-1, pp. 29-43, Jan. 2018
- Ge Chen, Le Yi Wang, Chen Chen, George Yin, Critical Connectivity and Fastest Convergence Rates of Distributed Consensus with **Switching** Topologies and Additive Noises, ***IEEE Transactions on Automatic Control***, Vol. 62-12, pp. 6152-6167, April 2017.
- K. Tran, G. Yin, L.Y. Wang, Hanqing Zhang, Singularly Perturbed Multi-scale **Switching** Diffusions, ***Dynamic Systems and Applications***, 25, pp. 153-174. 2016.

# Dealing with Random Delays

## Classical Systems

**One constant delay**

**Multiple constant delays**

**Time varying delay**

**Randomly time-varying and distributed delays**

**Interactive random delays with random noise and random network structures**

## Network System Reality

**Important and Promising New Methods:**
**Stochastic Functional Differential Systems and Functional Itô Formula**

# Theoretical Foundation (George Yin and his collaborators): Functional Itô Formula for Random Delays

## History:

- Extensive advance on stochastic delay equations, but no Itô formulas for distributed and integral-type delays.
- In "(2009) B. Dupire, Functional Itô's Calculus. Bloomberg Portfolio Research Paper No. 2009-04-FRONTIERS", Dupire extended the Itô formula to a functional setting using a path-wise functional derivative.
- Rigorous mathematical treatment: 2013 by R. Cont, D.-A. Fournie, "Functional Itô calculus and stochastic integral representation of martingales", Ann. Probab., 41, pp. 1109-1133, 2013
- This work created a new direction in studying stochastic functional equations.

## George Yin's work:

- We used this idea in establishing basic properties such as Feller properties, recurrence, positive recurrence, and ergodicity of switching diffusions.
    1. D.H. Nguyen, **G. Yin**, Modeling and analysis of switching diffusion systems: Past-dependent switching with a countable state space, ***SIAM J. Control Optim***. 54-5, pp. 2450-2477, 2016.
    2. D. Nguyen and **G. Yin**, Recurrence and ergodicity of switching diffusions with past-dependent switching having a countable state space, to appear in ***Potential Anal.***
- We applied this formula in studies of consensability in networked systems with distributed delays or more general functional-type uncertainties

  Xiaofeng Zong, George Yin, Le Yi Wang, Tao Li, Jifeng Zhang, Stability of stochastic functional differential systems using degenerate Lyapunov functionals and applications, ***Automatica***, Volume 91, pp. 197-207, May 2018.

## Functional space:

Let **D** be the space of cadlag (space of functions that are right continuous with left limits) functions $f : [-r,0] \to R^n$. For $\phi \in D$, we define the horizontal (time) and vertical (space) perturbations for $h \geq 0$ and $y \in R^n$

$$\phi_h(\mathrm{s}) = \begin{cases} \phi(s+h), \text{ if } s \in [-r,-h] \\ \phi(0), \text{ if } s \in [-h,0] \end{cases} ; \qquad \phi^y(\mathrm{s}) = \begin{cases} \phi(s), \text{ if } s \in [-r,0] \\ \phi(0)+y, \end{cases}$$

## Path-wise functional derivatives (with switching):

Let $V(\bullet,\bullet): \mathbf{D} \times Z_+ \to R$. The horizontal derivative at $(\phi,i)$ and vertical partial derivative of $V$ are defined as

$$V_t(\phi,i) = \lim_{h\to 0} \frac{V(\phi_h,i) - V(\phi,i)}{h}$$

$$V_x(\phi,i) = \lim_{h\to 0} \frac{V(\phi^{he_l},i) - V(\phi,i)}{h}$$

if these limits exist. $e_l$ is the standard unit vector in $R^n$ whose $l$-th component is 1 and other components are 0.

## Operators:

Let $\mathbf{F}$ be the family of function $V(\bullet,\bullet):\ \mathbf{D}\times Z_+ \to R$ satisfying that

(1) $V$ is continuous, that is, for any $\varepsilon>0$, $(\phi,i)\in\mathbf{D}\times Z_+$, there is a $\delta>0$ such that $|V(\phi,i)-V(\phi',i)|<\varepsilon$ as long as $\|\phi-\phi'\|<\delta$.

(2) The functions $V_t$, $V_x=(\partial_k V)$, and $V_{xx}=(\partial_{kl}V)$ exist and are continuous.

(3) $V_t$, $V_x$, and $V_{xx}$ are bounded in each $B_r:=\{(\phi,i):\ \|\phi\|\le r,\ i\le\ r\},\ r>0.$

For $V(\bullet,\bullet)\in\ \mathbf{F}$, we define the operator

$$\mathbf{L}V(\phi,\ i)=V_t(\phi,i)+V_x(\phi,i)b(\phi(0),i)+\frac{1}{2}Tr(V_{xx}(\phi,i)A(\phi(0),i))+\sum_{j=1,\,j\neq i}^{\infty}q_{ij}(\phi)[V(\phi,\ j)-V(\phi,\ i)]$$

## Functional Itô Formula:

For any bounded stopping time $\tau_1\le\tau_2$, we have the functional Itô formula:

$$E\ V(X_{\tau_2},\alpha(\tau_2))=V(X_{\tau_1},\alpha(\tau_1))+E\int_{\tau_1}^{\tau_2}\mathbf{L}V(X_s,\alpha(\tau_1))ds$$

if the involved expectations exist.

- Xiaofeng Zong, George Yin, Le Yi Wang, Tao Li, Ji-Feng Zhang, Stability of Stochastic Functional Differential Systems Using Degenerate Lyapunov Functionals and Applications, ***Automatica***, Volume 91, pp. 197-207, May 2018.
- Xiaofeng Zong, Tao Li, George Yin, Le Yi Wang, Ji-Feng Zhang, Stochastic Consentability of Linear Systems with Time Delays and Multiplicative Noises, ***IEEE Transactions on Automatic Control***, DOI: 10.1109/TAC.2017.2732823, July 2017.
- Lijian Xu, Le Yi Wang, George Yin, Hongwei Zhang, Impact of communication erasure channels on safety of highway vehicle platoons, ***IEEE Transactions on Intelligent Transportation Systems***, Vol. 16-3, pp. 1456-1468, June 2015.
- Fuke Wu, George Yin, Le Yi Wang, Razumikhin-type theorems on moment exponential stability of functional differential equations involving two-time-scale Markovian switching, ***Mathematical Control and Related Fields***, Vol. 5, No. 3, pp. 697-719, September 2015.
- Fuke Wu, George Yin, Le Yi Wang, Stability of a pure random delay system with two-time-scale Markovian switching, ***J. Differential Equations***, 253, 878–905, 2012.
- Wu, F.; Yin, G., Wang, L.Y.; Wang, Moment exponential stability of random delay systems with two-time-scale Markovian switching, ***Nonlinear Analysis Series B: Real World Applications***, Vol. 13, pp. 2476-2490, 2012.
- G. Yin, Le Yi Wang, Yu Sun: Stochastic Recursive Algorithms for Networked Systems with Delay and Random Switching: Multiscale Formulations and Asymptotic Properties, **SIAM J. *Multiscale Modeling & Simulation***, 9(3): 1087-1112, 2011.

# Challenging Open Question:
# Integration of Robust Control and Stochastic Systems

**Stochastic Approximation and Stochastic Functional Differential Equations:**

- State-Based Model
- Markov Decisions
- Stochastic Uncertainty (sampling, delays, noise, hybrid)

1. Combining **Deterministic** and **Stochastic** Systems
2. Combining **Model-based** and **Data-based** Methods
3. Combining **Robust** Design and **Learning/Adaptation**

**Distributed Robust Control (such as H-infinity):**

- Operator-Based Control
- Feedback
- Worst-Case Model Uncertainty

# Acknowledgement

## Collaborators on Networked Systems

- **Starting of Everything:** George Zames （老乔）
- **Everything Stochastic:** George Yin （殷刚）and many of his students and collaborators
- **Theoretical Aspects**: Lin Lin (藺林）， Jifeng Zhang （张纪峰）, Lei Guo （郭雷）, Hanfu Chen （陈翰馥）, Wei Xing Zheng （郑卫新）, Yanlong Zhao （赵延龙）, Fuke Wu （吴付科）, Chanying Li （李婵颖）, Ge Chen （陈鸽）, Jin Guo （郭金）, Biqiang Mu （牟必强）, Shu Liang （梁舒）, Wei Feng （冯巍）, Ping Zhao （赵平）, Xiaofeng Zong （宗小峰）, Tao Li （李韬）, Wenxiao Zhao (赵文虓）， and students
- **Battery and Renewable Energy**: Feng Lin （林峰）, Wen Chen （陈文）, Michael Polis, Jiuchun Jiang （姜久春）, Weige Zhang （张维戈）, Caiping Zhang （张彩萍）, Yan Bao （鲍谚）, Rui Xiong (熊瑞）, Quanqing Yu （于全庆）, and students
- **Autonomous Vehicles:** Hongwei Zhang （张洪伟）, Keqiang Li （李克强）, Shengbo Li （李升波), Feng Gao （高峰）, Lijian Xu （徐利剑）, and students
- **Smart Grid:** Caisheng Wang （王财胜）. Lei Ding （丁磊）, Yang Wang （汪洋）, Feixiong Chen (陈飞雄）, Minyou Chen (陈民铀）， Masoud Nazari, and students
- **Social Networks:** Hairong Dong （董海荣）, Bin Ning （宁斌）, Qianling Wang （王千龄）
- **Parallel Computing:** Cheng Zhong Xu (须成忠）, Aoda Fang （方敖大）, and students
- Many of my students and other collaborators

# Thank You!