基于有限信息数据的辨识与控制

张纪峰

中国科学院数学与系统科学研究院
系统科学研究所

对于一个给定的控制任务，
究竟需要多少信息量？

- 可否基于低精度数据给出高精度的辨识
- 可否基于低精度数据给出渐近最优的适应跟踪控制
- 可否基于1比特的通信数据实现趋同控制

对于一个给定的辨识或控制任务，究竟需要多少信息量？

内容提纲

- 问题背景
- 参数辨识
- 适应跟踪控制
- 多自主体系统控制
- 结束语

问题背景

状态反馈控制

被控系统

$$\dot{x} = Ax + Bu,$$
$$y = Cx + Du;$$

控制器

$$u = Kx$$

极点配置、
镇定控制、
LQ控制、
鲁棒控制、……

关键条件： 状态已知！

输出反馈控制

被控系统

$$\dot{x} = Ax + Bu,$$
$$y = Cx + Du;$$

控制器

$$u = Ky;$$

$$\begin{cases} \dot{\hat{x}} = Ny + Fu \\ u = K\hat{x} + Hy; \end{cases}$$

静态输出反馈控制、
动态输出反馈控制、
……

关键条件： 输出已知！

滤波、参数辨识、适应控制

- **系统模型：**

  * $\dot{x} = Ax + Bu$ —— 状态的量测噪声为高斯白噪声

  * $y_k = \sum_{\mathrm{i}=1}^{n} \mathrm{a}_{\mathrm{i}} y_{k-i} + \sum_{\mathrm{i}=1}^{n} \mathrm{b}_{\mathrm{i}} u_{k-i} + d_k = \theta^{\tau} \varphi_k + d_k$

- **辨识算法：LS、SG、LMS、KF、…**

- **关键条件：输出/状态是(基本)知道的！**

  (带有具有一定统计特性的量测噪声！)

最小二乘法

$$y_k = \sum_{\mathrm{i}=1}^{n} \mathrm{a}_{\mathrm{i}} y_{k-i} + \sum_{\mathrm{i}=1}^{n} \mathrm{b}_{\mathrm{i}} u_{k-i} + d_k = \theta^{\tau} \varphi_k + d_k$$

当y_k已知时，我们可用最小二乘法估计参数：

$$\hat{\theta}_k = \arg\min_{\theta \in R} \sum_{i=1}^{k} (y_i - \theta^{\tau} \varphi_{i-1})^2;$$

可用必然等价原则设计适应控制。

最小二乘法

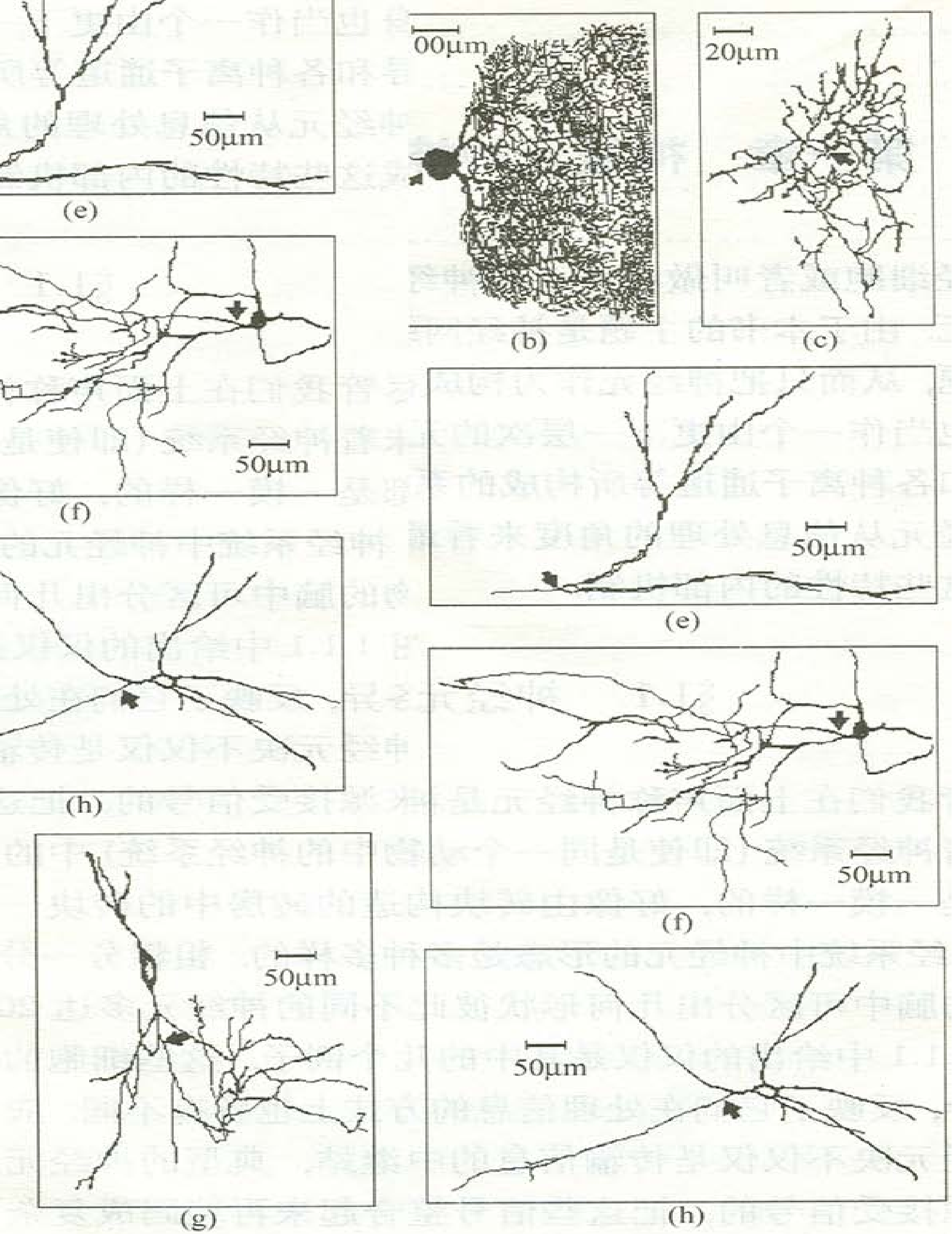
$$y_k=\sum_{\mathrm{i}=1}^{n}\mathrm{a_i}y_{k-i}+\sum_{\mathrm{i}=1}^{n}\mathrm{b_i}u_{k-i}+d_k=\theta^{\tau}\varphi_k+d_k$$

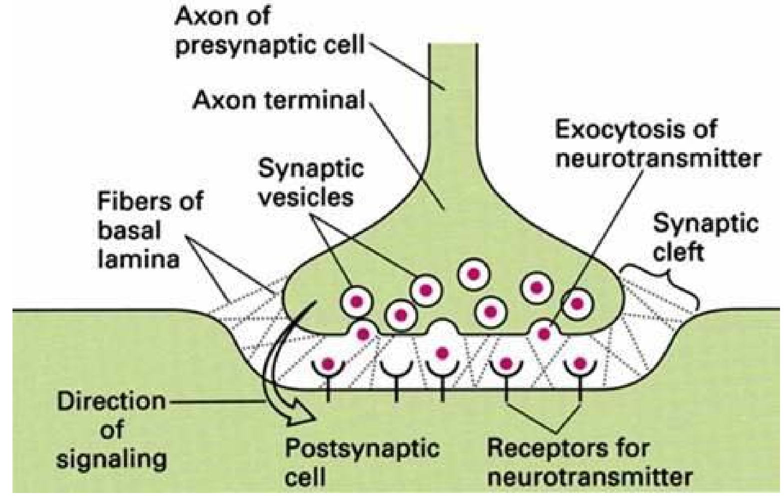
当y_k已知时，我们可用最小二乘法估计参数：

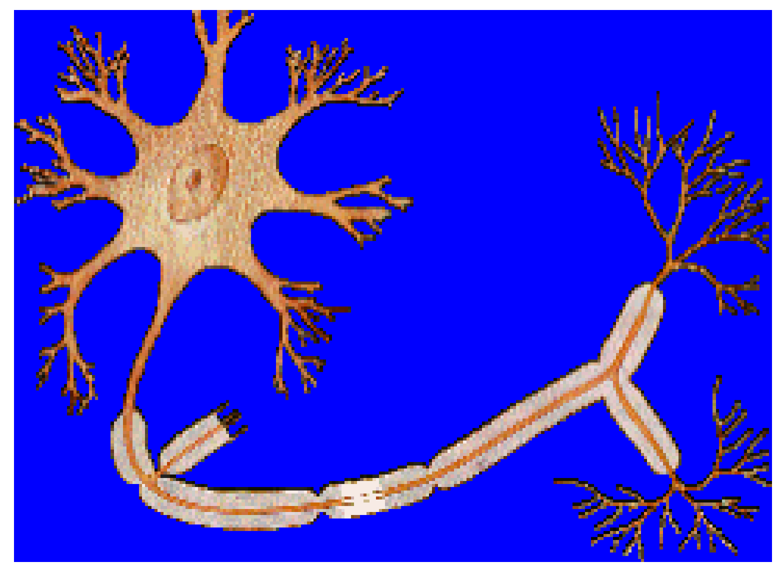
$$\hat{\theta}_k=\arg\min_{\theta\in R}\sum_{i=1}^{k}(y_i-\theta^{\tau}\varphi_{i-1})^2;$$

可用必然等价原则设计适应控制。

问题： 当系统的被控输出基本不知道时，可否也能设计出好的适应控制律?


00μm
20μm
50μm
(b)
(c)
(e)
(f)
(g)
(h)


Axon of
presynaptic cell
Axon terminal
Exocytosis of
neurotransmitter
Synaptic
vesicles
Fibers of
basal
lamina
Synaptic
cleft
Direction
of
signaling
Postsynaptic
cell
Receptors for
neurotransmitter



只有集值信息可用

难以拿到精度高的数据

传输带宽有限

捷赛自动烹饪锅——智能炒菜机具备煎鱼、炒菜、烹调、炸虾、爆料、焖饭、蒸馒头、煮粥、烙饼、炖肉、煲烫等烹饪技术的智能化，不会做饭的人也能轻松做出美味菜肴。　用精度高的温度敏感器件不合算!!

- 温度测量
- 品质控制

有限信息数据的辨识与控制

- 有时不需要精确的数据;
- 有时只有粗糙/集值数据可用;
- 有时精确数据难以得到;
- 有时使用精度高的测量元件不合算;
- 有时因带宽的限制只能得到量化数据;
- ……

科学问题：

可否基于低精度测量器件提供的数据给出高精度的模型和满意控制？

例如, 只能测到它与某个阈值的大小比较关系:

$$y_k \in (c_k, \infty) \quad \text{还是} \quad y_k \in (-\infty, c_k],$$

其中$y_k \in \mathrm{R}$是被控输出，$c_k \in \mathrm{R}$是阈值。

称为二集值输出系统

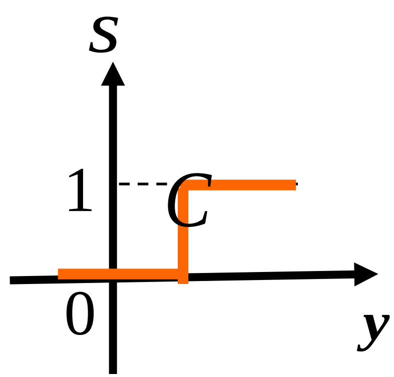
集值测量元件

■ 二集值测量元件

阈值

$$y_k \ge \mathrm{C} \text{ 或 } y_k < \mathrm{C}$$

$$q_k = \begin{cases} 1, & \text{if } y_k > C \\ 0, & \text{if } y_k \le C. \end{cases}$$

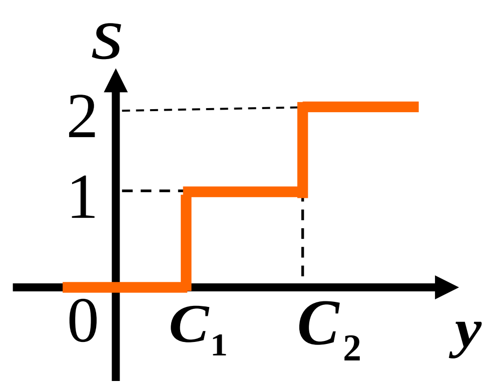


◆传感器网络
◆集成电路
◆网络通讯
◆机械系统
◆智能材料
◆汽车
◆化工
◆生物、…

■ 三集值测量元件

阈值

$$q_k = \begin{cases} 2, & \text{if } y_k > C_2; \\ 1, & \text{if } C_1 < y_k \le C_2; \\ 0, & \text{if } y_k \le C_1. \end{cases}$$



集值输出的特点

- 信息量少(字节有限)、非线性强:

A. 量测到的只是被控制信号与阈值的比较关系，不是被控制信号本身。

B. 与采样系统不同，在采样系统中，量测到的是采样时刻上系统输出的精确值。

C. 与通常的量化系统不同，在已有的量化系统的滤波、估计等工作中，要求量化误差充分小或者具有很好的统计特性(闭环条件)

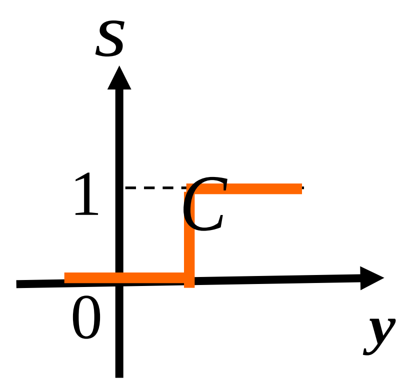
集值测量元件

- **二集值测量元件**

阈值

$$y_k \geq \mathrm{C} \text{ 或 } y_k < \mathrm{C}$$

$$q_k = \begin{cases} 1, \text{ if } y_k > C \\ 0, \text{ if } y_k \leq C. \end{cases}$$



y_k已知：$y_k = b\,u_k \Rightarrow b = y_1 / u_1$

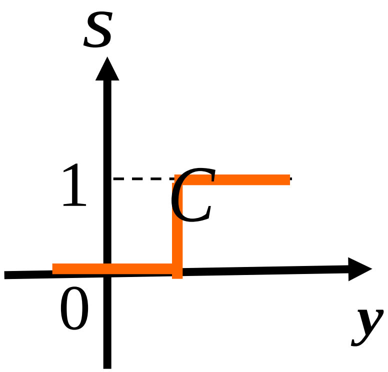
集值测量元件

■ 二集值测量元件

阈值

$$y_k \geq \mathrm{C} \text{ 或 } y_k < \mathrm{C}$$

$$q_k = \begin{cases} 1, \text{ if } y_k > C \\ 0, \text{ if } y_k \leq C. \end{cases}$$



y_k已知： $y_k = b\,u_k \Rightarrow b = y_1 / u_1$

y_k未知： $y_k = b\,u_k \not\Rightarrow b = y_1 / u_1$

参数辨识

关键科学问题

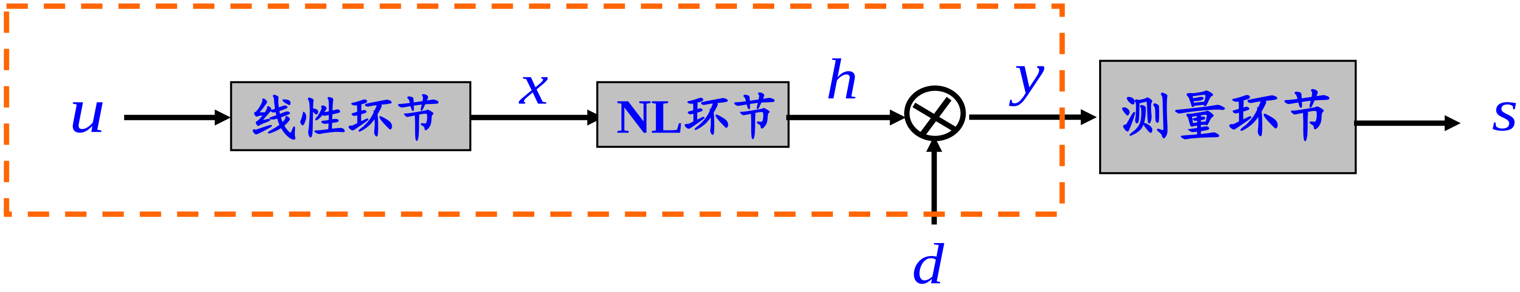
- 可辨识性
- 精度及收敛速度
- 输入信号设计
- 不确定性对辨识精度、收敛速度及计算复杂度的影响
- 阈值分布对辨识精度、收敛速度及计算复杂度的影响

系统模型

- **ARX：** $y(t) = \sum_{i=1}^{n} a_i y(t-i) + \sum_{i=0}^{n} b_i u(t-i) + d(t)$

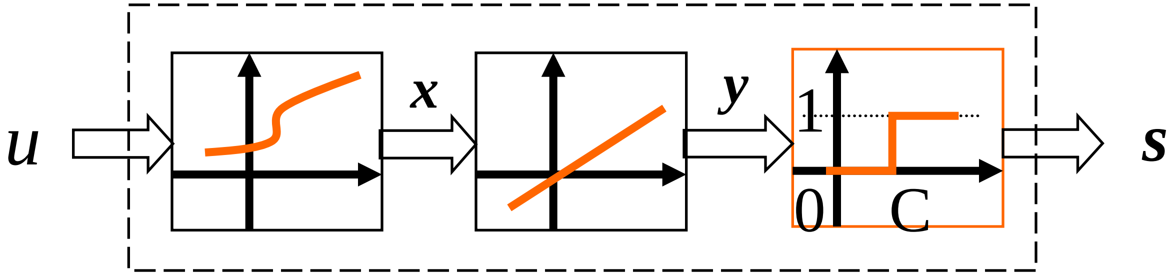
其中a_i、b_i是未知参数, $d(t)$是未知干扰

- **Wiener 系统:**



$$x(t) = \sum_{i=0}^{n-1} a_i u(t-i), \qquad y(t) = \sum_{i=0}^{m-1} x^i(t) b_i + d(t)$$

● **Hammerstein系统:**



$$\mathrm{NL}_1(\eta) \longrightarrow \mathbf{L}(\boldsymbol{\theta}) \longrightarrow \mathrm{NL}_2$$

模型:

$$\begin{cases} y(k) = \displaystyle\sum_{i=0}^{n-1} a_i x(k-i) + d(k), \\ x(k) = b_0 + \displaystyle\sum_{j=1}^{m} b_j u^j(k), \quad b_m = 1. \end{cases}$$

辨识算法

- 二分法 → $y(\mathrm{t}) = bu(\mathrm{t}) + d(\mathrm{t}),$

- 参数解耦法

$$y(\mathrm{t}) = \sum_{\mathrm{i}=1}^{n} \mathrm{b_i} u(t-i) + d(t)$$

- 持续激励
- 周期输入

$$y(\mathrm{t}) = \sum_{\mathrm{i}=1}^{n} \mathrm{a_i} y(t-i) + \sum_{\mathrm{i}=1}^{n} \mathrm{b_i} u(t-i) + d(t)$$

- 经验分布函数法：随机噪声情形

单参数情形

- **模型：** $y(\mathrm{t}) = bu(t)$

$b \in [\underline{b}(0), \overline{b}(0)],$

$0 < \underline{b}(0) < \overline{b}(0).$

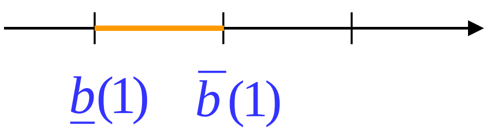
初值条件

- **输入：** $u(1) = \dfrac{2C}{\underline{b}(0) + \overline{b}(0)}$

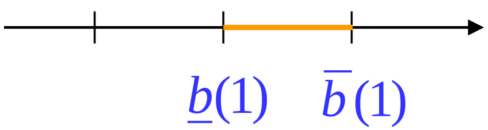
二集值

- **1-步辨识误差：**

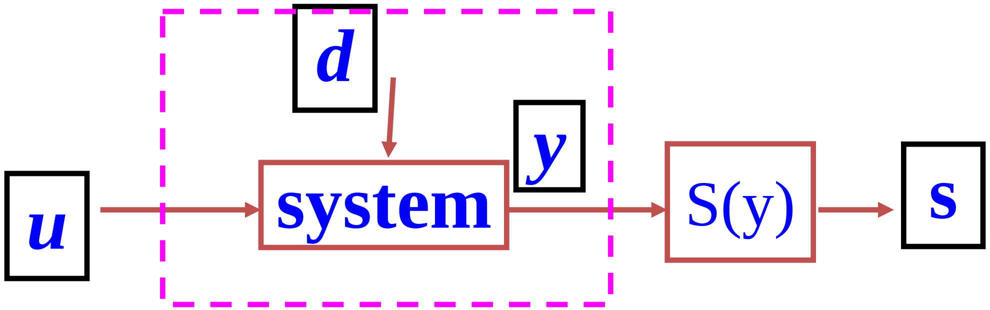
若$y(1) \le C$，则$b \le \overline{b}(1) = \dfrac{\underline{b}(0) + \overline{b}(0)}{2}$；



若$y(1) > C$，则$b > \underline{b}(1) = \dfrac{\underline{b}(0) + \overline{b}(0)}{2}$.



有限脉冲(FIR)系统



$$y(t) = \sum_{i=1}^{n} b_i u(t-i) + d(t)$$

$$s = S(x) = \begin{cases} 1, & \text{if } x > C_1; \\ j, & \text{if } C_j < x \le C_{j-1}, \\ & \quad j = 2,3,\ldots,m-1; \\ m, & \text{if } x \le C_{m-1}. \end{cases}$$

时间复杂度

定理1: **对FIR系统，设** $|d(t)| \le \delta,$

记 $\bar{b}(0) = \max\limits_{1\le i\le n} \bar{b}_i(0),\ \underline{b}(0) = \min\limits_{1\le i\le n} \underline{b}_i(0), \varepsilon(0) = \mathrm{Rad}(\Omega_0),$

$$\sigma = \frac{m(C_1 - C_2) + 2\delta}{2C_1 - C_2 - \delta}\bar{b}(0),\ \ \alpha_1 = \frac{(C_1 - C_{m-1} - 2\delta)\underline{b}(0)}{C_1 - \delta},$$

$\alpha_2 = (C_{m-1} + \delta)(C_1 + C_{m-1})^{-1},\ C_1 - C_2 = \ldots = C_{m-2} - C_{m-1},$

则 $\forall \varepsilon \in (\sigma, \varepsilon(0))$, 有

$$\|d\|_\infty \le \delta$$

$$N(\varepsilon) \le \frac{\ell_n \ln \dfrac{\alpha_1 + \varepsilon n^{-1/2}}{(1-\alpha_1)\varepsilon(0)}}{\ln \alpha_2}.$$

适应跟踪控制

集值输出系统的适应控制

考虑如下的单参数系统:

$$y_k = \theta u_{k-1} + d_k,$$

量测输出为:

$$s_k = I_{[y_k \le c_k]} = \begin{cases} 1, \text{ if } y_k \le C, \\ 0, \text{ otherwise}, \end{cases}$$

其中u_k、d_k、$C \in \mathrm{R}$分别是输入、噪声和阈值;

$y_k \in \mathrm{R}$是系统的被控输出，目标是极小化:

$$J_k = E(y_k - y_k^*)^2.$$

目的：在参数未知时，给出适应控制律

假设条件

(A1) 已知正实数$\overline{\theta}$和$\underline{\theta}$, $0<\underline{\theta}<\overline{\theta}<\infty$, 使得$|\theta|\in[\underline{\theta},\overline{\theta}]$.

(A2) 已知正实数$\overline{y}^*$和$\underline{y}^*$, $0<\underline{y}^*<\overline{y}^*<\infty$, 使得

$|y_k^*|\in[\underline{y}^*,\overline{y}^*]$.

(A3) $\{d_k, k\geq 1\}$是零均值、有限方差iid随机序列，d_1的分布函数$F(x)$已知, 密度函数$f(x)$关于$x=0$对称且在$(0,\infty)$上单调递减，其支撑集包含$[-T,T]$，即$[-T,T]\subseteq\{x: f(x)\neq 0\}$。这里$T=|C|+3\overline{\theta}\overline{y}^*/\underline{\theta}$。

辨识算法

记 $\Theta = [-\bar{\theta}, \bar{\theta}]$, 在$\Theta$定义投影算子:

$$\Pi_{\Theta}(x) = \arg\min_{z\in\Theta} |x - z|,$$

则θ的估计由下式给出

$$\hat{\theta}_k = \Pi_{\Theta}\{\hat{\theta}_{k-1} + \frac{\beta P_{k-1} u_k}{1 + P_{k-1} u_k^2}[F(C - \hat{\theta}_{k-1} u_k) - s_k]\},$$

$$P_k = P_{k-1} - \frac{\gamma P_{k-1}^2 u_k^2}{1 + P_{k-1} u_k^2},$$

其中$|\hat{\theta}_0| \in [\underline{\theta}, \bar{\theta}]$, $P_0 > 0$, $\beta > 0$, $\gamma \in (0,1]$, $F(\cdot)$是噪声d_k的分布函数.

辨识算法

$$\hat{\theta}_k = \Pi_{\Theta}\{\hat{\theta}_{k-1} + \frac{\beta P_{k-1} u_k}{1 + P_{k-1} u_k^2}[F(C - \hat{\theta}_{k-1} u_k) - s_k]\},$$

$$P_k = P_{k-1} - \frac{\gamma P_{k-1}^2 u_k^2}{1 + P_{k-1} u_k^2},$$

$$s_k = I_{[y_k \le C]}$$

$F(\cdot)$ 是 d_k 的分布函数

常规LS ➡

$$\begin{cases} \hat{\theta}_k = \hat{\theta}_{k-1} + \dfrac{P_{k-1} u_k}{1 + P_{k-1} u_k^2}(y_k - \hat{\theta}_{k-1} u_k), \\ P_k = P_{k-1} - \dfrac{\gamma P_{k-1}^2 u_k^2}{1 + P_{k-1} u_k^2}. \end{cases}$$

适应控制律

$$u_k = \frac{y_k^*}{\hat{\theta}_{k-1}} I_{[\underline{\theta} \le |\hat{\theta}_{k-1}| \le \overline{\theta}]} + \frac{y_k^*}{\underline{\theta}} \left(I_{[0 < \hat{\theta}_{k-1} < \underline{\theta}]} - I_{[-\underline{\theta} < \hat{\theta}_{k-1} \le 0]} \right) \tag{3}$$

定理2：在假设(A1)-(A3)下，有

◆参数估计

- 以概率1收敛到真值：$\lim_{k\to\infty}\hat{\theta}_k=\theta, \text{a.s.}$
- 渐近有效(达到**Cramer-Rao**下界),
- 估计误差的均方收敛速度为

$$E\|\hat{\theta}_k-\theta\|^2=O(k^{-1/2})$$

◆适应跟踪渐近最优

$$E\|y_k-y_k^*\|^2=\delta+O(k^{-1/2})$$

数值算例-1

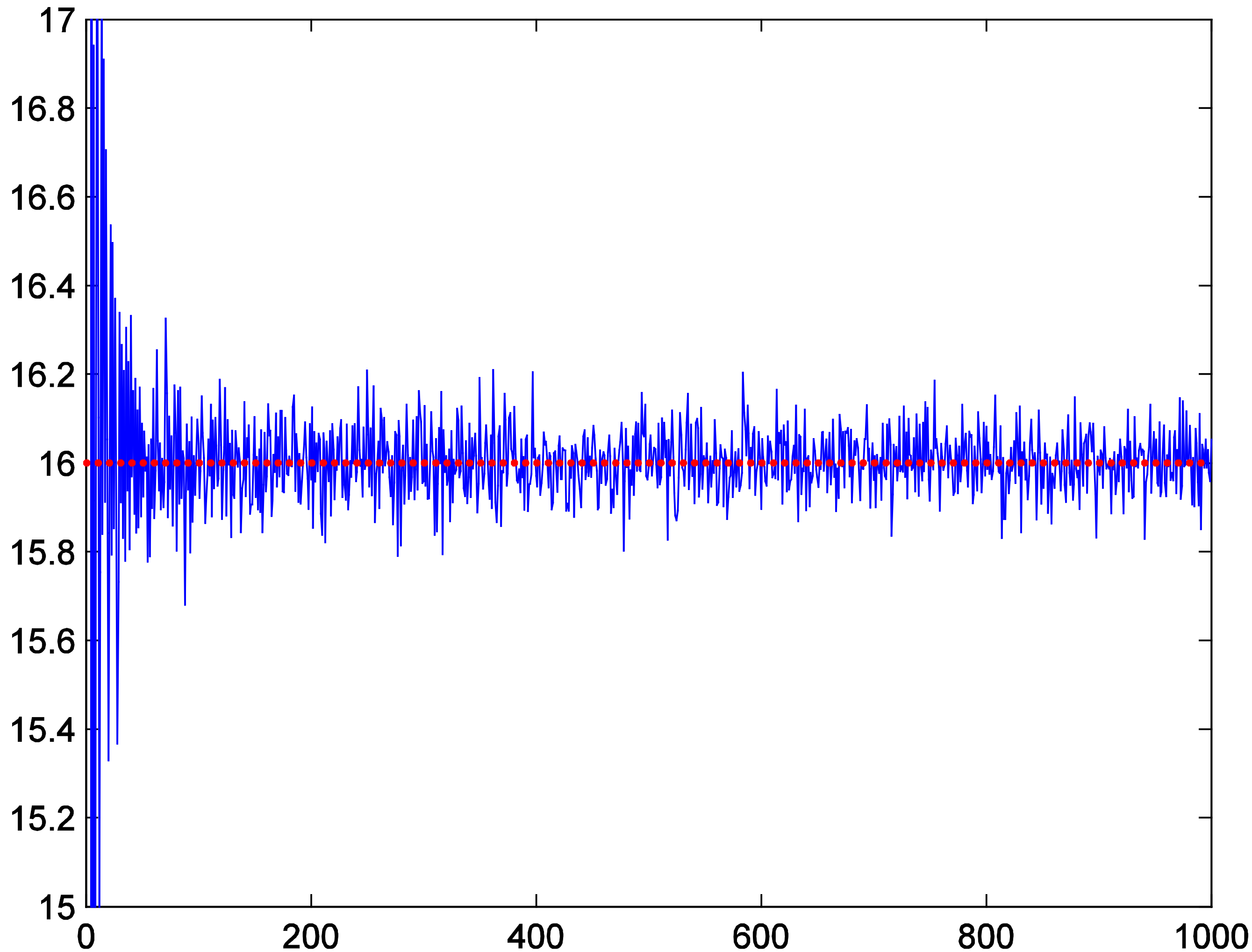
◆模型： $y_k = 6u_{k-1} + d_k$

◆阈值： **C=2**

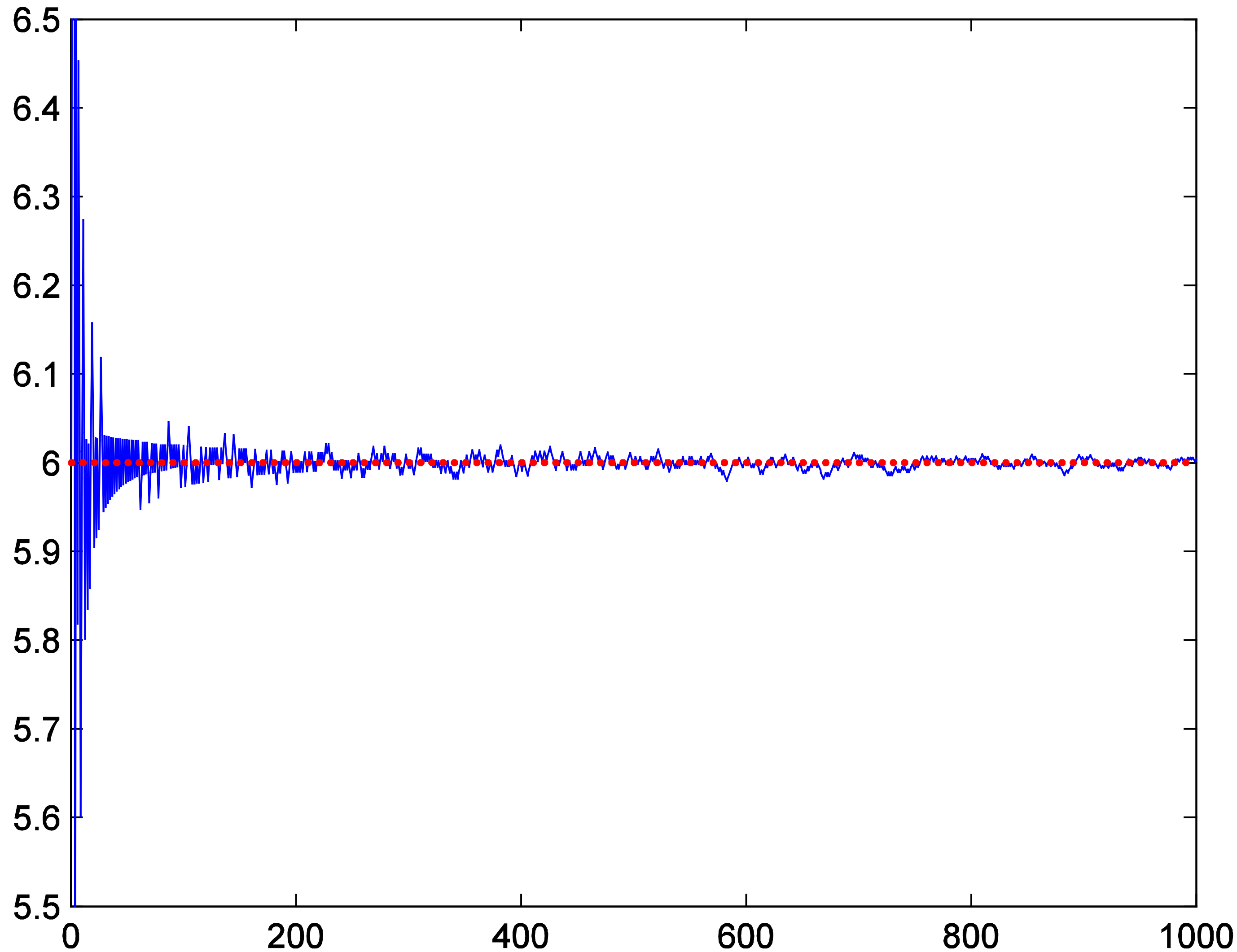
◆初始值:

$$\beta = 1,\quad \gamma = 0.5,\quad y_k^* \equiv 16,\quad F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{0.2\pi}} e^{-50\mu^2} d\mu,$$

$$\theta = 6,\quad \underline{\theta} = 1,\quad \overline{\theta} = 15,\quad \hat{\theta}_0 = P_0 = 1.$$



Tracking performance of y_k (solid) to $y^* \equiv 16$ (dashed)



Estimation convergence of $\hat{\theta}_k$ (solid) to the true parameter $\theta = 6$

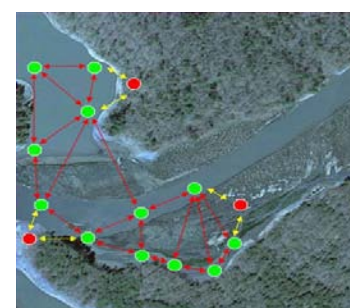
多自主体系统控制

应用背景

- 队形控制
- 分布式滤波
- 多传感器信息融合
- 分布式计算

基于远程传感器网络的多自主体系统的协调控制

分布式趋同控制



能量的优化利用

传输1比特
消耗的能量

≈

进行1000-3000次
运算消耗的能量

Shnayder 等
2004

极小化传输字节
以节省能量

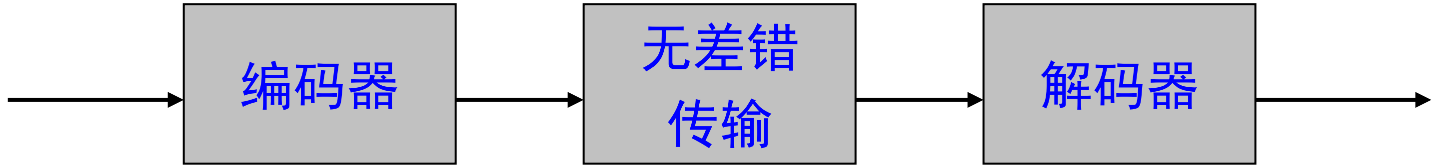
量化传输

- ➢ Kashyap et al. Automatica 2007
- ➢ Carli et al. ECC 2007
- ➢ Carli et al. IFAC 2008
- ➢ Kar & Moura arXiv:0712 2007

- **特点：**

✓ 量化层数为无穷

✓ 非零稳态误差



关键科学问题

- 可否基于有限的数据传输带宽得到以指数速度趋同？

- 如果可以，为实现趋同，至少需要多大带宽？

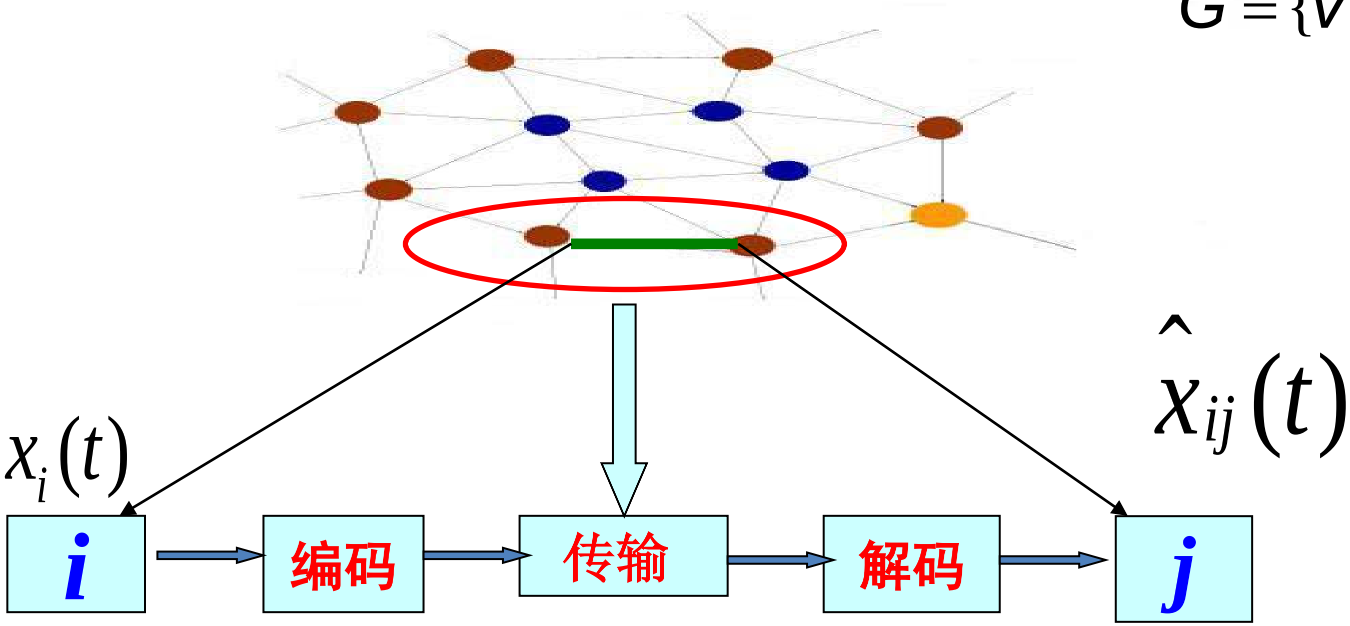
- 带宽与趋同速度的具体关系如何？

要实现趋同，每对邻居每次通讯需要交换多少比特的信息？

系统模型

$$x_i(t+1) = x_i(t) + hu_i(t),\ t = 0,1,\ldots,\ i = 1,2\ldots,N$$

$G = \{V, E, A\}$



状态是实值的，但每次只能传输有限bit的数据

- **难点**
  - 量化层数有限：非线性、量化误差无界
  - 量化误差：可能会导致状态发散

- **关键问题：**
  - 编码器、解码器、趋同控制协议的设计
    - ✓ 镇定系统
    - ✓ 消除量化误差的影响

模型、编码器、量化器

- 动态网络模型 $x_i(t+1)=x_i(t)+hu_i(t),\ t=0,1,\ldots,\ i=1,2\ldots,N$
- 无向网络 $G=\{V,E,A\}$
- 编码器φ_j $\begin{cases}\xi_j(0)=0,\\ \xi_j(t)=g(t-1)\Delta_j(t)+\xi_j(t-1)\\ \Delta_j(t)=q[(g(t-1))^{-1}(x_j(t)-\xi_j(t-1))]\end{cases}$

尺度函数

新息量化

- 量化器

$$q(y)=\begin{cases}0, & -1/2<y<1/2\\ i, & (2i-1)/2\le y<(2i+1)/2\\ K, & y\ge(2K-1)/2\\ -(q(-y)), & y\le -1/2\end{cases}$$

(2K+1) 层

$\lceil \log_2(2K)\rceil$ **bits**

趋同协议

- 解码器ψ_{ji}:

$$\begin{cases} \hat{x}_{ji}(0) = 0, \\ \hat{x}_{ji}(t) = g(t-1)\Delta_j(t) + \hat{x}_{ji}(t-1),\ t = 1, 2, \ldots \end{cases}$$

- 分布式协议:

$$u_i(t) = \sum_{j \in N_i} a_{ij}(\hat{x}_{ji}(t) - \xi_i(t)),\ t = 0, 1, \ldots,\ i = 1, 2, \ldots, N$$

$$\begin{aligned} u_i(t) = & \sum_{j \in N_i} a_{ij}(x_j(t) - x_i(t)) \\ & - \sum_{j \in N_i} a_{ij}(x_j(t) - \hat{x}_{ji}(t)) \\ & + \sum_{j \in N_i} a_{ji}(x_i(t) - \hat{x}_{ij}(t)) \end{aligned}$$

误差补偿项

趋同协议

- 解码器ψ_{ji} :

$$\begin{cases} \hat{x}_{ji}(0) = 0, \\ \hat{x}_{ji}(t) = g(t-1)\Delta_j(t) + \hat{x}_{ji}(t-1),\ t = 1, 2, \ldots \end{cases}$$

- 分布式协议:

$$u_i(t) = \sum_{j \in N_i} a_{ij}(\hat{x}_{ji}(t) - \xi_i(t)),\ t = 0, 1, \ldots,\ i = 1, 2, \ldots, N$$

$$\begin{aligned} u_i(t) = & \sum_{j \in N_i} a_{ij}(x_j(t) - x_i(t)) \\ & - \sum_{j \in N_i} a_{ij}(x_j(t) - \hat{x}_{ji}(t)) \\ & + \sum_{j \in N_i} a_{ji}(x_i(t) - \hat{x}_{ij}(t)) \end{aligned}$$

$$\frac{1}{N}\sum_{i=1}^{N} x_i(t+1) = \frac{1}{N}\sum_{i=1}^{N} x_i(t)$$

关键参数

控制增益： h

量化层数： $2K+1$

尺度函数： $g(t)$

趋同速度

主要假设

(B1) **G 是连通图**

(B2) $\max\limits_{1\le i\le N} |x_i(0)| \le C_x, \quad \max\limits_{1\le i\le N} |\delta_i(0)| \le C_\delta$

记 $\rho_h = \max\limits_{2\le i\le N} |1 - h\lambda_i|$

G的拉氏阵的第i个特征值

$$\delta(t) = \left[x_1(t) - \frac{1}{N}\sum_{j=0}^{N} x_j(0), \cdots, x_N(t) - \frac{1}{N}\sum_{j=0}^{N} x_j(0) \right]^T$$

定理3:

若(B1)-(B2)成立，则对任意的 $K \geq 1$, $h \in (0, 2/\lambda_N)$，

$$\Omega_K = \left\{ (h,\gamma) \mid 0 < h < \frac{2}{\lambda_2 + \lambda_N},\ \rho_h < \gamma < 1,\ M_1(h,\gamma) < K + \frac{1}{2} \right\}$$

是非空的；对任意的 $(h,\gamma) \in \Omega_K$，若取 $g(t) = g_0 \gamma^t$,

则有

$$\lim_{t\to\infty} x_i(t) = \frac{1}{N}\sum_{j=1}^{N} x_j(0),\ i = 1, 2, \ldots, N$$

$$\|\delta(t)\| = O(\gamma^t), \quad t \to \infty,$$

其中 $$M_1(h,\gamma) = \frac{\sqrt{N} h^2 \lambda_N^2}{2\gamma(\gamma - \rho_h)} + \frac{1 + h\lambda_N}{2\gamma}$$

趋同速度

若G 是连通的, 则对任给的 $K \geq 1$, 有

$$\lim_{N\to\infty}\frac{\inf_{(h,\gamma)\in\Omega_K}\gamma}{\exp\left\{-\frac{KQ_N^2}{2\sqrt{N}}\right\}}=1$$

其中

$$Q_N = \frac{\lambda_2}{\lambda_N}$$

趋同速度

若G 是连通的, 则对任给的 $K \geq 1$，有

$$\lim_{N\to\infty} \frac{\inf_{(h,\gamma)\in\Omega_K} \gamma}{\exp\left\{-\frac{KQ_N^2}{2\sqrt{N}}\right\}} = 1$$

其中

$$Q_N = \frac{\lambda_2}{\lambda_N}$$

Synchronizability

Barahona等 PRL 2002

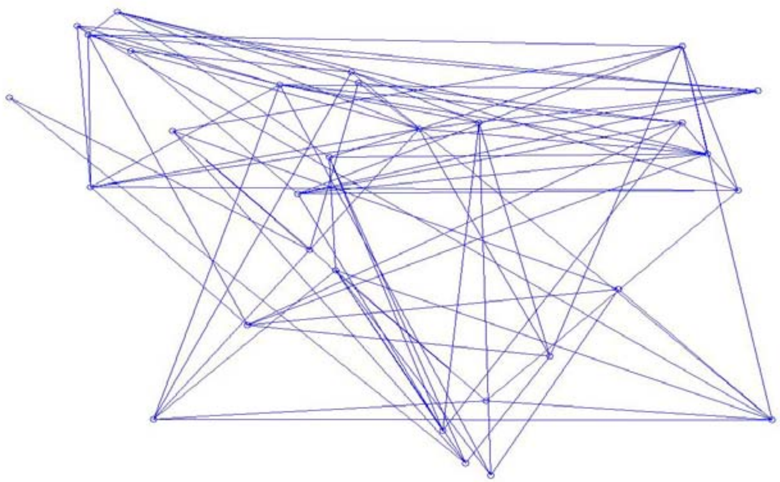
Donetti 等 PRL 2005

趋同速度

收敛速度随量化层数和同步参数的增加而提高，但随个体数的增加而减小

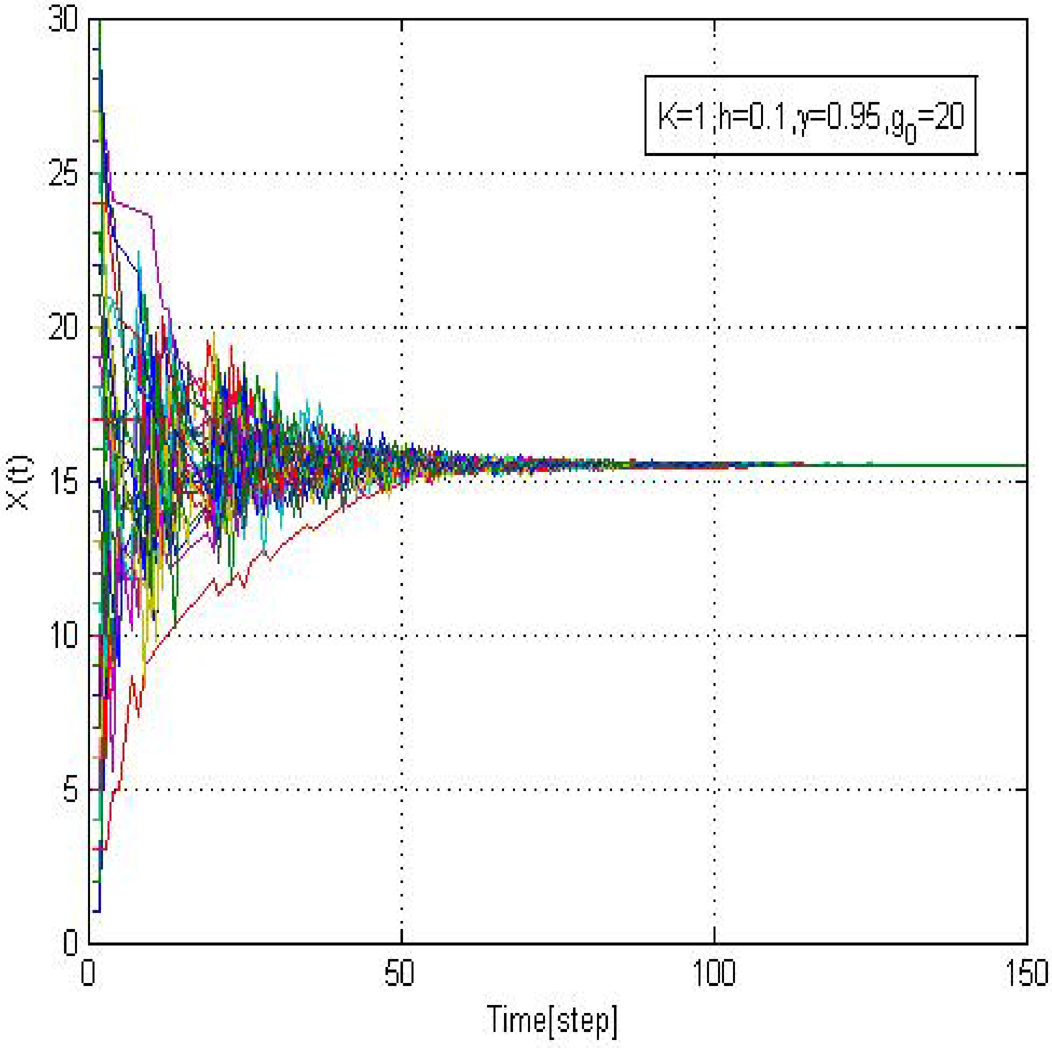
$$\inf_{(h,\gamma)\in\Omega_K} \gamma \approx \exp\left\{-\frac{KQ_N^2}{2\sqrt{N}}\right\}$$

数值算例-2

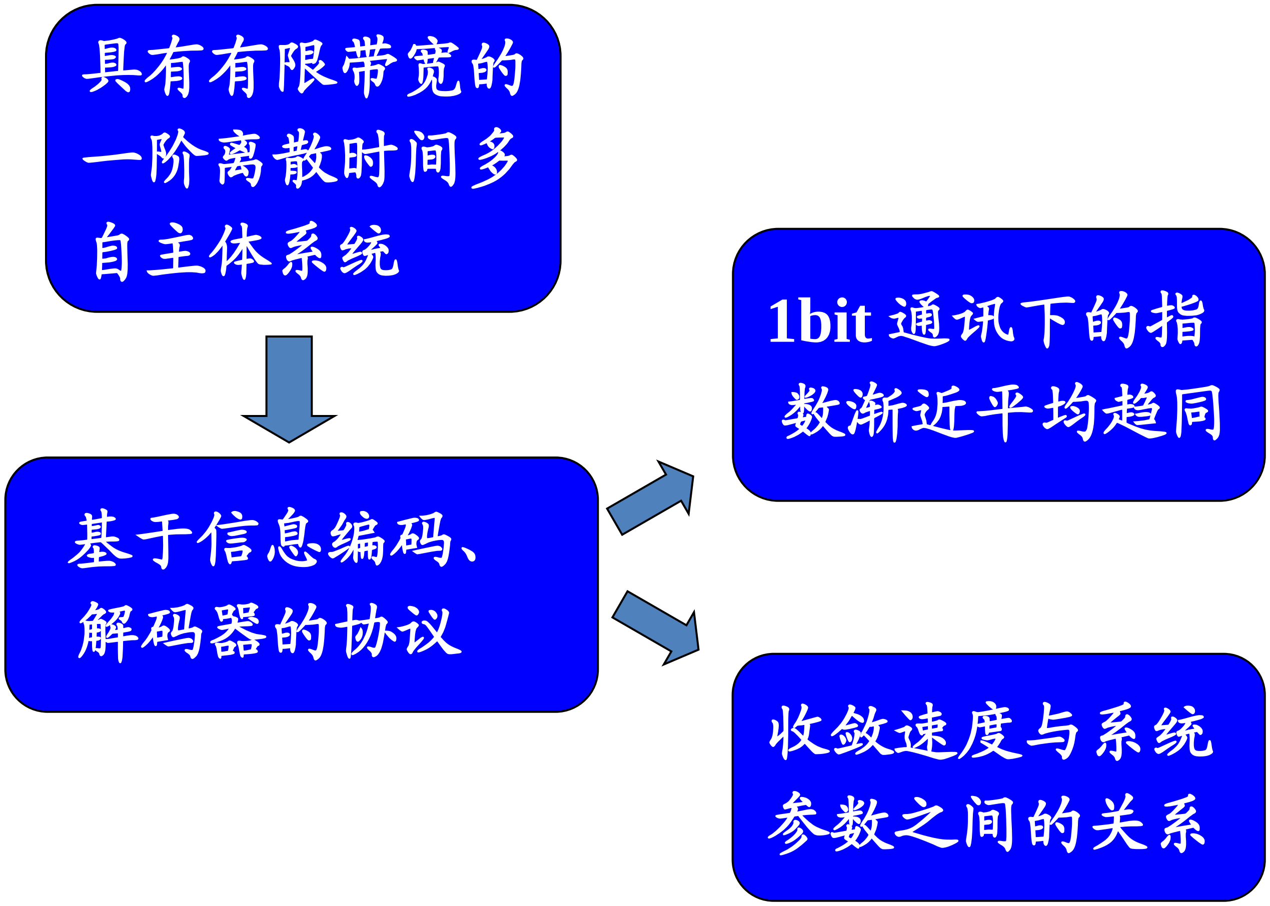


具有30个节点的网络

- 边的选取: $P\{(i,j)\in E\}=0.2$
- 初始状态: $x_i(0)=i,\ i=1,2,\dots,30$



1比特量化器，收敛因子0.95


具有有限带宽的一阶离散时间多自主体系统
基于信息编码、解码器的协议
1bit通讯下的指数渐近平均趋同
收敛速度与系统参数之间的关系

结束语

- 可基于低精度数据给出高精度的辨识
- 可基于低精度数据给出渐近最优的适应跟踪控制
- 可基于1比特的通信数据实现趋同控制

- 可基于低精度数据给出高精度的辨识
- 可基于低精度数据给出渐近最优的适应跟踪控制
- 可基于1比特的通信数据实现趋同控制

对于一个给定的辨识或控制任务，究竟需要多少信息量?

任务、约束与复杂性

- 任务

  建模、辨识还是控制

- 约束

  机理约束、测量条件、成本约束、时间/速度要求、精度要求、通讯及信息处理能力、……

- 复杂性

  计算复杂度、实现复杂度、理论分析的复杂度等

F(任务、约束、复杂性)=?

谢谢