arXiv:2303.03113v2 [nlin.AO] 7 Mar 2023

定量　　基于粗采样时间序列的因果关系分析

X. San Liang(日

期：2023年3月8日)摘要

基于信息流的定量因果分析因源自第一性原理、形式简洁且计算效率高，已广泛应用于不同学科。迄今为止，其估计算法基于微分动力系统；然而，这可能会给粗采样时间序列带来问题。本文表明，对于线性系统，至少在定性层面上这并无大碍；但对于高度非线性系统，采样频率降低时，偏差会显著增大。本文针对这一问题提供了部分解决方案，说明当统计量足够时，如何确保对粗采样序列进行因果关系分析仍能忠实反映因果性。本文得到了一个明确而简洁的公式，其中只涉及样本协方差。该公式已成功应用于由一对耦合的 R¨ossler 振子组成的系统。尤其值得注意的是，当两个振子近乎同步时，该公式仍能取得成功。

PACS 分类号：
关键词：定量因果关系；信息流；粗采样时间序列；同步；R¨ossler 系统；李群

一、引言

因果关系分析是科学研究中的一个重要问题。尽管传统上它被表述为数据科学、计算机科学等学科中的统计问题，但近来研究发现，在信息流／传递框架下，它是“物理学中的一种真实概念，可以从头推导出来”[1]。近期已针对一般系统开展了全面研究，并得到了明确的闭式公式；参见 [2] 和 [1]。这些公式已通过贝克变换、H´énon 映射等基准系统得到验证，并已成功应用于多个学科的实际问题，例如全球气候变化（如 [3], [4], [5]）、动态气象学（如[6]）、陆气相互作用（如[7]）、数据驱动预测（如 [8], [9]）、近壁湍流（如 [10]）、神经科学（如 [11], [12]）、金融分析（如 [13]¸[14]）、量子信息（如 [15]）等。

为开展本研究，我们首先简要介绍微分动力系统框架下的相关理论。（该理论也适用于离散时间映射，参见 Liang (2016)。）设

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}, t) + \mathbf{B}(\mathbf{x}, t)\dot{\mathbf{w}}, \tag{1}$$

为一个关于 x $=(x_1, ..., x_d)$ 的 d 维连续时间随机系统（我们不区分随机变量与确定性变量的记号），其中 F $=(F_1, ..., F_d)$ 可以是关于 x 和 t 的任意非线性可微函数，w 是白噪声向量，而 B $=(b_{ij})$ 是扰动幅度矩阵，也可以是关于 x 和 t 的任意可微函数。Liang (2016)[1] 证明，从 x_j 到x_i 的信息流率（单位时间内的奈特数）为

$$\begin{aligned} T_{j\to i} &= -E\left[\frac{1}{\rho_i}\int_{\mathbb{R}^{d-2}}\frac{\partial(F_i\rho_{\not ij})}{\partial x_i}d\mathbf{x}_{\not ij}\right] + \frac{1}{2}E\left[\frac{1}{\rho_i}\int_{\mathbb{R}^{d-2}}\frac{\partial^2(g_{ii}\rho_{\not ij})}{\partial x_i^2}d\mathbf{x}_{\not ij}\right], \\ &= -\int_{\mathbb{R}^d}\rho_{j|i}(x_j|x_i)\frac{\partial(F_i\rho_{\not ij})}{\partial x_i}d\mathbf{x} + \frac{1}{2}\int_{\mathbb{R}^d}\rho_{j|i}(x_j|x_i)\frac{\partial^2(g_{ii}\rho_{\not ij})}{\partial x_i^2}d\mathbf{x}, \end{aligned} \tag{2}$$

其中， $d\mathbf{x}_{\not ij}$ 表示 $dx_1...dx_{i-1}dx_{i+1}...dx_{j-1}dx_{j+1}...dx_n$， E 表示数学期望，$g_{ii} = \sum_{k=1}^{n} b_{ik}b_{ik}$， $\rho_i = \rho_i(x_i)$ 是x_i的边缘概率密度函数（pdf），$\rho_{j|i}$ 是给定 x_i时 x_j 的概率密度函数（pdf），而 $\rho_{\not ij} = \int_{\mathbb{R}} \rho(\mathbf{x})dx_j$。基于信息流的因果推断算法如下：若 $T_{j\to i} = 0$，则 x_j 对x_i没有因果作用；否则，二者存在因果关系，且其绝对值衡量从 x_j 到 x_i的因果作用强度。这由一项称为“零因果原理”的性质所保证。另一项性质涉及坐标变换下的不变性，表明所得到的信息流（IF）在本质上是一种内禀属性[16]。Liang (2016)[1] 还证明，对于线性模型，即 F(x t) $=$ Ax，且 (1) 中的 A $=(a_{ij})$ 和 B $=(b_{ij})$ 为常数矩阵，则

$$T_{j\to i} = a_{ij}\frac{\sigma_{ij}}{\sigma_{ii}}$$

其中， σ_{ij} 是 x_i 和 x_j的总体协方差。由此可知，在线性意义上，因果关系意味着相关性，但反之则不成立。明确地说，这条推论为自乔治·贝克莱（1710）[17]以来关于因果关系与相关性的争论作出了定论。

对于仅有 d 个时间序列 x_1、x_2、.....、x_d的情况，它们之间的定量因果关系，即信息流，可以通过最大似然估计来估计（参见 [18] 和 [19]）。在线性系统具有加性噪声这一假设下，式（2）针对 $T_{j\to i}$ 的最大似然估计量（mle）为[19]

$$\hat{T}_{j\to i} = \frac{1}{\det \mathbf{C}} \cdot \sum_{\nu=1}^{d} \Delta_{j\nu} C_{\nu,di} \cdot \frac{C_{ij}}{C_{ii}}, \tag{3}$$

其中，C_{ij} 是 x_i 与 x_j之间的样本协方差，Δ_{ij} 是矩阵 $\mathsf{C} = (C_{ij})$ 的余子式，而$C_{i,\ dj}$ 是 x_i 与利用欧拉向前差分格式从 x_j 导出的序列之间的样本协方差：$\dot{x}_{j,\ n} = (x_{j,\ n+k} - x_{j,\ n})/(k\Delta t)$，其中 $k \geq 1$ 为某个整数。式（3）的形式相当简洁，只涉及常见统计量，即样本协方差。这个直观明了的公式使因果关系分析变得十分简单且计算高效，而原本这项分析会很复杂。不过请注意，式（3）不能取代式（2）；它只是后者的最大似然估计量（mle）。可以对这些估计量进行统计显著性检验，方法是借助费 Fisher 信息矩阵。详情参见 Liang (2014)[18] 和 Liang (2021)[19] 。

该形式体系最初是在微分系统的框架下建立的；换言之，它适用于无穷小时间增量。（基于离散映射的形式体系也已由 Liang (2016)[1], 建立，但尚未据此进行估计。）人们自然会问，它是否适用于粗采样时间序列。事实上，由于观测有限，给定序列采用粗采样并不少见。正如下一节将要看到的，如果采样间隔较大，这可能会给非线性系统带来问题。因此，本文尝试在线性框架下探讨这一问题。下文首先考察式（3）对来自线性系统和高度非线性系统的序列是否适用（第二节），并采用多种采样间隔。第三节提出一种新方法，随后用它重新进行第二节中的因果推断。第五节讨论一些尚待解决的问题。

II. 粗采样时间序列的问题

A. 线性系统的时间序列

我们首先检验（3）式的适用性，考察其在采样间隔增大时的表现。为此，我们采用一个研究充分的线性系统，其信息流率已通过半解析方法求得。以下是 [18]中的验证示例：

$$\frac{dx_1}{dt} = -x_1 + 0.5x_2 + 0.1\dot{w}_1 \tag{4a}$$

$$\frac{dx_2}{dt} = -x_2 + 0.1\dot{w}_2 \tag{4b}$$

其中，$\dot{w}_i$、$i = 1$和 2 是相互独立的白噪声。已有研究表明，单位时间的信息流率为 $T_{2\to 1} \to 0.11$ as $t \to \infty$，且对所有 t均有 $T_{1\to 2} = 0$ ，准确地反映了从 x_2 到 x_1的单向因果关系。现在，我们采用与 Liang (2014) 中相同的样本路径，[21] 以低频率对序列重新采样，从而得到新的序列。图 1 展示了部分样本路径，三角形标出了采样点。

不同采样间隔下计算出的 IF 如下：

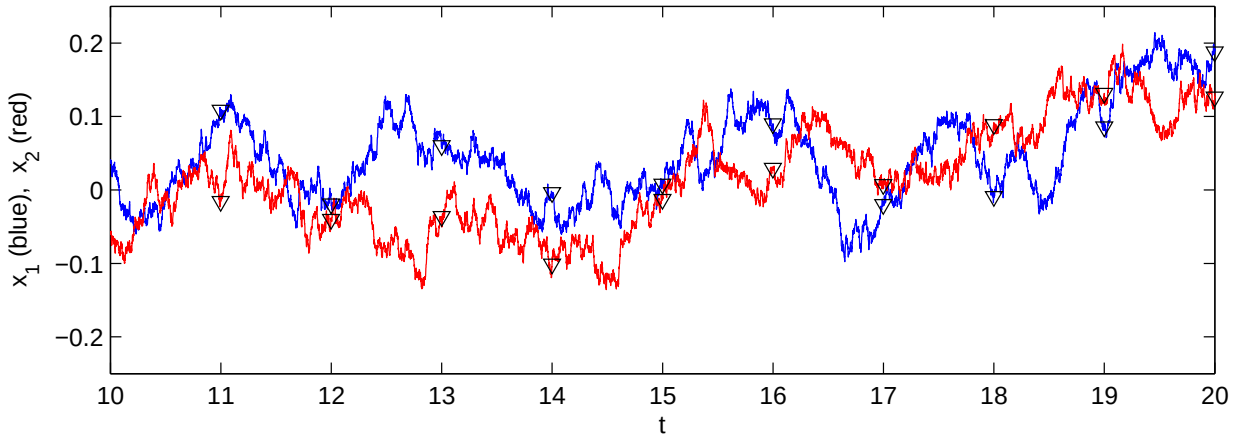


图 1：由方程 (4) 生成的样本路径片段，时间步长为 $\Delta t = 0.001$。时间范围为 0 到 100。由于我们只关注 $t \to \infty$ 时的极限信息流，因此舍弃前 5000 个时间步，即 $t \leq 5$ 对应的时间步。

采样间隔（点数）	1	10	50	100	300	500
$\hat{T}_{2\to1}$	0.110±0.051	0.113±0.051	0.099±0.050	0.090±0.048	0.055±0.036	0.055±0.034
$\hat{T}_{1\to2}$	-0.002±0.056	-0.001±0.055	-0.015±0.054	-0.011±0.053	-0.008±0.038	-0.015±0.034

此外，还计算了置信水平为 90%（显著性水平为 0.1）时的置信区间。ˆ首先，此处所有采样间隔对应的估计量 $T_{2\to1}$ 均显著异于零，而ˆ相反的那些估计量 $T_{1\to2}$在 90% 的置信水平下并不显著。因此，即使采样频率很低（采样间隔较大），因果性在定性意义上仍得到了忠实恢复。（事实上，即使采样间隔为 SI=1000 ，结果仍然正确；由于 SI¿500 时样本量过小，导致统计量不足，因此我们不考虑超出 SI=500 的情况。）

由于这个例子实际上有一个半解析解（$T_{2\to1} \approx 0.11$, $T_{1\to2} = 0$），因此我们可以对计算结果做更多分析。总体而言， $T_{1\to2}$ 的结果看起来令人满意。对于 $T_{2\to1}$，当采样间隔为 ≤ 100 时，结果相当准确。超过 100 后，结果就不再准确。

B. 同步混沌振子的时间序列

以下示例来自 Palus 等人（2018）研究的同步问题。该系统由两个 R¨ossler 振子组成，即 x $= (x_1, x_2, x_3)$ 和 y $= (y_1, y_2, y_3)$，其中

$$\frac{dx_1}{dt} = -\omega_1 x_2 - x_3, \tag{5a}$$

$$\frac{dx_2}{dt} = \omega_1 x_1 + 0.15 x_2, \tag{5b}$$

$$\frac{dx_3}{dt} = 0.2 + x_3(x_1 - 10), \tag{5c}$$

是主系统，而

$$\frac{dy_1}{dt} = -\omega_2 y_2 - y_3 + \varepsilon(x_1 - y_1), \tag{6a}$$

$$\frac{dy_2}{dt} = \omega_2 y_1 + 0.15 y_2, \tag{6b}$$

$$\frac{dy_3}{dt} = 0.2 + y_3(y_1 - 10), \tag{6c}$$

是受驱动的。按照 Palus 等人（2018）[20], 选择 $\omega_1 = 1.015$和 $\omega_2 = 0.985$。使用龙格-库塔格式并选择时间步长 $\Delta t = 0.001$，耦合六维系统可以针对不同的 ε 得到相当精确的求解。图 2 绘出了耦合强度为 $\varepsilon = 0.11$ （上图）和 $\varepsilon = 0.15$ （下图）时 x_1和 y_1 的解。如后者所示，如果 $\varepsilon \geq 0.15$，两个子系统将变得同步。再次，我们

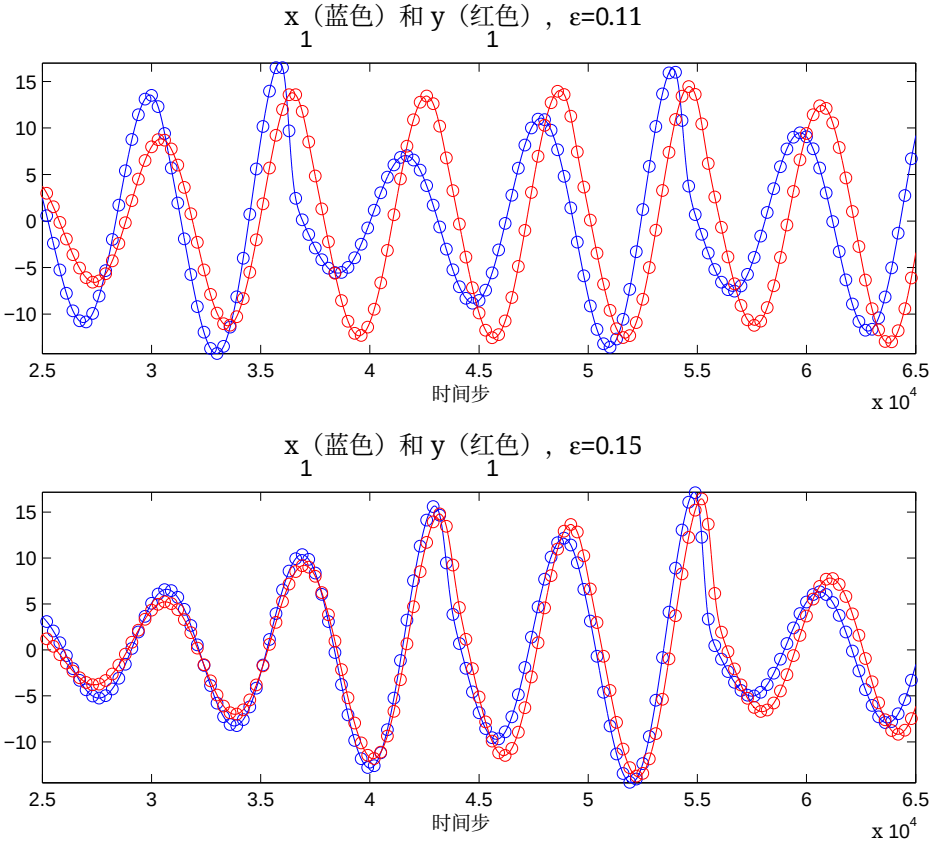


图 2：耦合 R¨ossler 系统中 x_1 和 y_1 的部分时间序列，分别对应 $\varepsilon = 0.11$ （上）和$\varepsilon = 0.15$ （下）。圆圈表示采样点（此处每 300 个步长一个，对应每个周期中的 18 个点）。随着 $\varepsilon \geq 0.15$，两个振子变得同步。

选择研究 $t \in [0, 100]$ 的问题（总共 10^5 个时间步）。

对于每个 ε ，我们生成六个长度为 10^5 步的时间序列，并根据公式 (3) 评估 IF（选取 $k = 1$ ）。随后得到 IF 随 ε 变化的函数，并绘制在图 3a 中。图中准确地表明，x 是主系统，y 是从系统。一个引人注目的

观察结果是，即使两个振子几乎同步为 $\varepsilon > 0.15$，这种因果推断仍然有效，展现了这种经过严格表述的因果关系分析的强大能力。

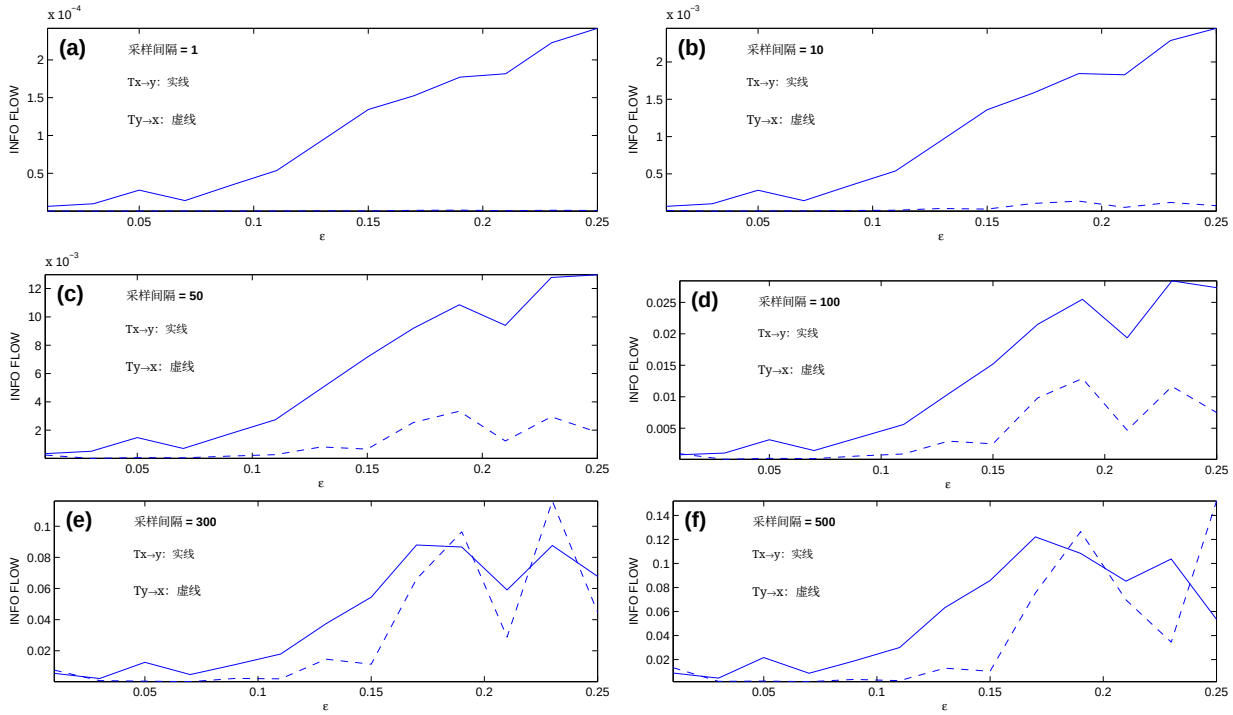


图 3：两个 R¨ossler 振子之间的绝对信息流率 $\left|\hat{T}_{x\to y}\right|$ 和 $\left|\hat{T}_{y\to x}\right|$（虚线）随耦合系数 ε 变化的情况，采样间隔各不相同。单位为每单位时间的奈特。根据预设的因果性，虚线应与横坐标轴重合。

我们每隔 SI 步对序列进行一次子采样，SI=10、50、100、300、500，并使用相同方案重新计算。所得信息流率分别见图 3b、c、d、e、f。根据预设的因果性，虚线应为零线。显然，当 SI$\leq$ 10 时，因果推断效果良好。当 SI$\geq$ 50 时，计算得到的信息流会产生偏差，且偏差随 SI 增大而显著增加。如果我们关注 $\varepsilon <= 0.15$，即系统不同步的情况（见 Palus 等，2018），那么当 SI$\leq$50 时，因果推断仍然表现良好。如果将同步的情况也纳入考虑（$\varepsilon > 0.15$），那么在 $50 \leq$SI ≤ 100 的情况下，推断会有很大偏差；而对于 SI 超过 300 的情况（相当于每个周期约 20 个采样点），推断则不再正确。

III. 逐步接近部分解

如上所示，如果时间序列的采样频率较低，所得针对非线性序列的线性信息流可能存在偏差。事实上，在非线性很强的情况下，线性假设总是很容易受到质疑。尽管理论上这并不是问题（如某一定理所证明，因果性得到了保证），但我们同意，在开发出完全非线性算法之前，这将一直是一个问题。我们想在此展示的是，改进的空间有多大。到目前为止，[18], 中记载并在 [19], 中进一步介绍的算法是

基于伯恩斯坦–欧拉差分方案，当然，由于采用一阶差分，该方案非常初步。如果一个时间序列被粗采样，误差可能会很大。

Liang（2008）[2] 提出的一个定理指出，如果噪声在公式（1）中是加性的，即 B 是一个常数矩阵，那么噪声本身不会出现在 $T_{j\to i}$ 的公式中。因此，在加性噪声假设下，我们可以在确定性系统的框架内估计信息流速率。在这种情况下，请注意，无论 Δt 有多大，实际上都可以在区间 $[t, t+\Delta t]$ 内求解该线性方程组。这为解决低采样频率问题提供了线索。

考虑

$$\frac{d\mathbf{x}}{dt} = \mathbf{f} + \mathbf{A}\mathbf{x}, \tag{7}$$

其中，$\mathbf{A} = (a_{ij})$ 是一个 $d \times d$ 矩阵。由于时间序列始终可以通过去除线性趋势进行预处理，且已有证明表明，这种去除不会改变信息流速率，因此我们假设 $\mathbf{f} = \mathbf{0}$。在这种情况下，在区间 $[t, t+\Delta t]$ 上，我们实际上有一个映射 $\Phi:\ \mathbb{R}^d \to \mathbb{R}^d$，它通过传播算子将状态 $\mathbf{x}(t)$ 映射到 $t+\Delta t$ 时的状态 $\mathbf{x}(t+\Delta t)$：

$$\Phi = e^{\mathbf{A}\Delta t} = e^{\begin{bmatrix} a_{11} & \dots & a_{1d} \\ \vdots & \ddots & \vdots \\ a_{d1} & \dots & a_{dd} \end{bmatrix}\Delta t} \equiv \begin{bmatrix} \alpha_{11} & \dots & \alpha_{1d} \\ \vdots & \ddots & \vdots \\ \alpha_{d1} & \dots & \alpha_{dd} \end{bmatrix}. \tag{8}$$

这并不容易估　计 a_{ij}，但通过观察这一关系，可以很容易地改为估计 α_{ij}

$$\begin{bmatrix} \alpha_{11} & \dots & \alpha_{1d} \\ \vdots & \ddots & \vdots \\ \alpha_{d1} & \dots & \alpha_{dd} \end{bmatrix} \mathbf{x}(n) = \mathbf{x}(n+1), \qquad n = 0, 1, 2, ..., N. \tag{9}$$

将其写成矩阵形式为

$$\begin{bmatrix} x_1(0) & \dots & x_d(0) \\ \vdots & \ddots & \vdots \\ x_1(N-1) & \dots & x_d(N-1) \end{bmatrix} \begin{bmatrix} \alpha_{i1} \\ \vdots \\ \alpha_{id} \end{bmatrix} = \begin{bmatrix} x_i(1) \\ \vdots \\ x_i(N) \end{bmatrix}$$

对于 $i = 1,.., d$。对代数方程组的所有行求均值，并从每一行中减去该均值，我们得到

$$\begin{bmatrix} x_1(0) - \bar{x}_1 & \dots & x_d(0) - \bar{x}_d \\ \vdots & \ddots & \vdots \\ x_1(N-1) - \bar{x}_1 & \dots & x_d(N-1) - \bar{x}_d \end{bmatrix} \begin{bmatrix} \alpha_{i1} \\ \vdots \\ \alpha_{id} \end{bmatrix} = \begin{bmatrix} x_i(1) - \bar{x}_{i+} \\ \vdots \\ x_i(N) - \bar{x}_{i+} \end{bmatrix},$$

其中 $\bar{x}_i = \frac{1}{N}\sum_{n=0}^{N-1} x_i(n)$、 $\bar{x}_{i+} = \frac{1}{N}\sum_{n=1}^{N} x_i(n)$，也就是说，级数 $\{x_{i+}(n)\}$ 是级数 $\{x_i(n)\}$ 向前推进一步得到的。令 i 遍历 $\{1, 2, ..., d\}$。我们得到以下 d 超定方程组：

$$\begin{bmatrix} x_1(0) - \bar{x}_1 & \dots & x_d(0) - \bar{x}_d \\ \vdots & \ddots & \vdots \\ x_1(N-1) - \bar{x}_1 & \dots & x_d(N-1) - \bar{x}_d \end{bmatrix} \begin{bmatrix} \alpha_{11} & \dots & \alpha_{d1} \\ \vdots & \ddots & \vdots \\ \alpha_{1d} & \vdots & \alpha_{dd} \end{bmatrix} = \begin{bmatrix} x_1(1) - \bar{x}_{1+} & \vdots & x_d(1) - \bar{x}_{d+} \\ \vdots & \ddots & \vdots \\ x_1(N) - \bar{x}_{1+} & \vdots & x_d(N) - \bar{x}_{d+} \end{bmatrix} \tag{10}$$

记 $\mathbf{\Lambda}$ 为矩阵（α_{ij}），则上述方程组中的未知数矩阵为$\mathbf{\Lambda}^T$。左乘

$$\begin{bmatrix} x_1(0)-\bar{x}_1 & \dots & x_d(0)-\bar{x}_d \\ \vdots & \ddots & \vdots \\ x_1(N-1)-\bar{x}_1 & \dots & x_d(N-1)-\bar{x}_d \end{bmatrix}^T$$

在两边得到 d $d \times d$ 方程组：

$$\mathbf{C}\mathbf{\Lambda}^T = \tilde{\mathbf{C}}, \tag{11}$$

˜其中 C $= (C_{ij})$ 是 x 的样本协方差矩阵，而 C $= (C_{i,j+})$ 和 $C_{i,j+}$ 是 x_i 与 x_{j+}之间的样本协方差，即 x_j 向前推进一个时间步长。超定方程组 (10) 的最小二乘解即为 (11) 的解：

$$\mathbf{\Lambda}^T = \mathbf{C}^{-1}\tilde{\mathbf{C}},$$

因此

$$\mathbf{\Lambda} = (\mathbf{C}^{-1}\tilde{\mathbf{C}})^T = \tilde{\mathbf{C}}^T\mathbf{C}^{-1}.$$

因此，A 的估计量为

$$\hat{\mathbf{A}} = \frac{1}{\Delta t}\log\left(\tilde{\mathbf{C}}^T\mathbf{C}^{-1}\right). \tag{12}$$

（若存在奇异性，需谨慎处理。无关的虚部也应舍弃。）

得到 A 后，也就得到了系数（a_{ij}），我们将 a_{ij} 代入整个部分

$$\frac{1}{\det \mathbf{C}}\sum_{k=1}^{d}\Delta_{jk}C_{k,di}$$

ˆ 在公式 (3) 中，即将 a_{ij} 乘以 C_{ij}/C_{ii} 以得到所需结果 $T_{j\to i}$。如果我们用 $[\mathrm{A}]_{ij}$ 表示提取矩阵 A 的 $(i,j)^{th}$ 项，则有

$$\hat{T}_{j\to i} = \frac{1}{\Delta t}\left[\log(\tilde{\mathbf{C}}^T\mathbf{C}^{-1})\right]_{ij}\cdot\frac{C_{ij}}{C_{ii}}. \tag{13}$$

（此处的 log 指矩阵对数。在 MATLAB 中，相应的函数是 logm。）

IV。粗采样序列问题再探讨

如上所示，对于由线性系统生成的序列，信息流的估计在定性上是可靠的。不过，我们仍希望了解新方案能否改善结果。下方展示了重新计算得到的估计值。由于这一情形的结果相当准确（每单位时间 ≈ 0.11 奈特），可以看出，此处所有采样间隔下的结果都足够准确。

采样间隔	1	10	50	100	300	500
$\hat{T}_{2\to1}$	0.114	0.118	0.109	0.106	0.082	0.098
$\hat{T}_{1\to2}$	0.007	0.008	−0.007	−0.002	−0.002	−0.015

这一用于估计的新方案尤其适用于非线性情形。对于一对 R¨ossler 振子，计算结果绘于图 4。与图 3 相比，现有方案的表现已有显著改善。对于采样间隔为 SI$\leq$ 100 的情形（图 4a-d），在所考察的所有耦合强度 ε 下，结果都相当准确（包括同步和非同步情形）。对于采样间隔为 SI=300 的情形，即采样频率为每周期 20 个点时，非同步情形下的单向因果关系得到了准确恢复（$\varepsilon \leq 0.15$）。但超过这一范围 $\varepsilon > 0.15$ 后，推断便会失败。尤其是当采样间隔为 SI=500 时（图 4f），结果甚至比图 3f 所示传统方案对应的结果更差。当然，这可能是由于样本量过小，导致矩阵对数出现奇异性。

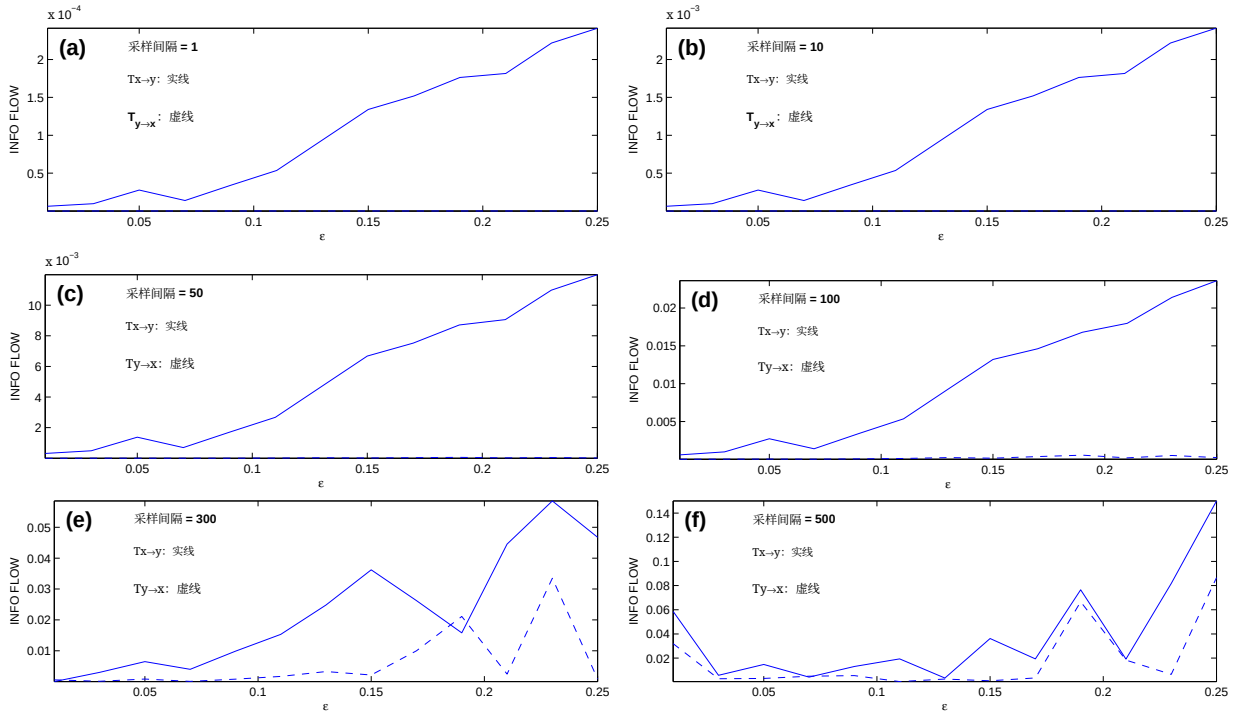


FIG. 4: 与图 3 相同，但信息流率采用新方案计算 .

V. 讨论

信息流（IF）的最大似然估计量（公式（3））为因果推断提供了一种非常简便的方法。从理论上说，它基于线性假设；但在实践中，对于高度非线性系统生成的序列，它已取得巨大成功；无论如何，分段时间线性化已证明是逼近原本非线性系统的一种有效渐近方法。实际上，序列可能经过粗采样，时间分辨率也可能较低。因此便产生了一个问题：这一形式体系在理论上以无穷小时间

增量为基础。在这种情况下，正如我们所展示的，它对线性系统仍具有定性上的适用性；但对于由两个 R¨ossler 振子组成的高度非线性系统，随着采样频率降低，偏差会越来越显著。

为解决这一问题并提供部分解决方案，研究提出了一种新方案。正如 [2], 所证明的，信息流（IF）具有一个很好的性质：加性噪声不会改变 IF 流的形式。因此，直接估计 IF 而不考虑随机性是合理的。我们没有采用 Euler-Bernstein 差分法通过微分方程进行估计，而是选择考虑有限时间区间上的积分形式，即估计李群元素。为此，将原公式 (3) 改写如下，以便于查阅：

$$\hat{T}_{j\to i} = \frac{1}{\det \mathbf{C}} \cdot \sum_{\nu=1}^{d} \Delta_{j\nu} C_{\nu,di} \cdot \frac{C_{ij}}{C_{ii}},$$

将其替换为 (13)：

$$\hat{T}_{j\to i} = \frac{1}{\Delta t} \left[\log(\tilde{\mathbf{C}}^T \mathbf{C}^{-1})\right]_{ij} \cdot \frac{C_{ij}}{C_{ii}},$$

其中 $\tilde{\mathbf{C}} = (C_{i,\ j+})$，且 $C_{i,\ j+}$ 是 x_i 与 x_{j+}之间的样本协方差，即 x_j 向前推进一个时间步长后的值。请注意，此处的 log 表示矩阵对数；在 MATLAB 中，对应的函数是 logm。由此可见，即使采样间隔较大（采样频率较低），耦合混沌振子系统中预设的因果性仍得到了相当准确的再现。

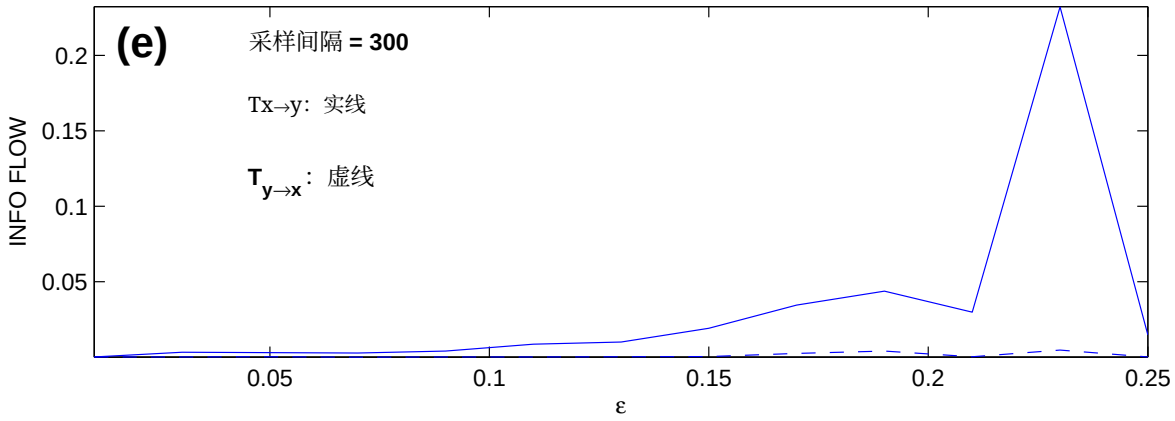


图 5：与图4e相同，但协方差是利用 x_i 相对于模型结果的残差估算的，而不是利用 x_i 本身。此处，采样间隔 SI=300 大致对应于每个周期采样 18 个点的采样频率。

上述方法仍有很大的改进空间。例如，商式 $\frac{\sigma_{ij}}{\sigma_{ii}}$ 中协方差的估计，是用样本协方差替代总体协方差，而样本则由时间序列构成。对于满足遍历假设的随机系统，这种做法令人满意；但对于确定性混沌，这种做法可能并不理想，例如此处的 R¨ossler 振子就是如此。原因显而易见：图 2 中序列的时间平均值为零，但可以想象，所有可能路径的系综平均值绝非零；

相反，它应该是时间的函数（就像序列本身一样），并且可能接近渐近线性系统解。因此，将线性系统解视为均值更为合理。基于此，我们尝试通过用 $x - \bar{x}$ 的协方差替代 x 的协方差来改进估计，其中 $\bar{x}$ 表示由此得到的线性系统解；借此，我们得到了关于 SI =300 的另一项因果推断结果；所得 IF 绘于图5中。可以看出，结果相当准确，正如预期，与图4e形成鲜明对比。

不过，我们并不声称已经解决了这个问题。我们在此想要展示的是，线性框架内还有多大的改进空间。事实上，在非线性很强的情况下，线性假设总是很容易受到诟病。从理论上说，这并不是问题（正如定理所证明的那样，因果性是有保证的；参见 [1] 及其他参考文献），但人们认为，在完全非线性算法开发出来之前，这个问题仍将持续存在。

代码可用性

代码已提供，并将在 www.ncoads.org/article/show/67.aspx 上更新。

致谢

非常感谢 Milan Paluˇs 提出这个问题，它促成了本项研究。

---

[1] Liang, X.S.（2016）。信息流与因果性作为严格的先验概念。《物理评论E》，94，052201。[2] Liang, X.S. 随机动力系统中的信息流。《物理评论E》2008，78，031113。[3] Stips, A.; Macias, D.; Coughlan, C.; Garcia-Gorriz, E.; Liang, X.S. 关于 CO_2 与全球温度之间的因果结构。《科学报告》2016，6，21691。[4]Vannitsem, S.; Dalaiden, Q.; Goosse, H.（2019）检验动力依赖——南极洲地表质量平衡的应用。《地球物理研究快报》，doi:10.1029/2019GL084329。[5] Doquier, D., S. Vannitsem, F. Ragone, K. Wyser, and X.S. Liang（2022）。基于信息传递速率研究北极海冰及其潜在驱动因素之间的因果联系。《地球物理研究快报》。DOI: 10.1029/2021GL095892。[6] Liang, X.S.（2019）。风暴模型中纬度大气的跨尺度因果关系与信息流研究。《熵》，21，149。[7] Hagan, D.F.T.; Wang, G.; Liang, X.S.; Dolman, H.A.J.（2019）基于Liang-Kleeman信息流的时变因果形式体系：用于分析非平稳气候系统中的定向相互作用。《气候杂志》，32，7521–7537。[8] Bai, C., R. Zhang, and Coauthors（2018）。通过识别气旋—气候相互作用中的因果因素，预测西北太平洋的热带气旋生成。《大气与海洋技术杂志》，35，247-259。

[9] Liang, X.S., F. Xu, Y. Rong, R. Zhang, X. Tang 和 F. Zhang（2021）。El Ni˜no Modoki 可提前十年以上进行较为准确的预测。《科学报告》，11，17860。[10] Liang, X.S., A. Lozano-Dur´an（2016）。充分发展近壁湍流因果结构的初步研究。2016年暑期项目论文集，233-242。湍流研究中心，斯坦福大学，加利福尼亚州，美国。[11] Hristopulos, D.T., A. Babul, S. Babul, L.R. Brucar, N. Virji-Babul（2019）。运动相关脑震荡青少年静息态中的信息流受损。《人类神经科学前沿》，13，419。doi:10.3389/fnhum.2019.00419。[12] Cong, J. 等（2023）。通过 Liang信息流分析揭示自闭症谱系障碍中默认模式网络因果连接模式的改变。《人脑图谱》，DOI:10.1002/hbm.26209。[13] Lu, X., K. Liu 等（2022）。CO_2 排放配额价格与清洁能源指数之间突发事件的动态因果关系。《环境科学与污染研究》。https://doi.org/10.1007/s11356-022-21316-5。[14] Lu, X.F., K. Liu, H. Cui, J. Li（2023）。国际原油期货市场与中国股指期货市场之间因果关系的研究。《中国管理科学》。http://doi.org/10.16381/j.cnki.issn1003-207x.2021.1128。[15] Yi, B., S. Bose（2022）。作为因果量化指标的量子Liang信息流。《物理评论快报》129，020501。[16] Liang, X.S.（2018）。关于相对熵的信息流。《混沌》28，075311。[17] Berkeley, G.《人类知识原理论》；Aaron Rhames：都柏林，爱尔兰，1710。[18] Liang, X.S.（2014）。揭示时间序列之间的因果关系。《物理评论E》90，052150。[19] Liang, X.S.（2021）。归一化多元时间序列因果分析与因果图重构。《熵》23，679。[20] Paluˇs、M.、A. Krakovsk´、J. Jakubfk 和 M. Chvostekov´（2018）。因果性、动力系统与时间之箭。《混沌》28，075307。[21] 请注意，由于伪随机数生成器的影响，使用常规差分方案生成的样本路径可能不尽如人意。要判断所得样本路径是否正确，可以检查相应的协方差；通过求解确定性常微分方程可以相当准确地得到这些协方差。此处生成样本路径所用的数据可从 http://www.ncoads.org/article/show/68.aspx 下载，文件名为 PRE 2014.dat。