## 基于阵列多通道数据的非线性特征参数提取\*

李惟嘉1)2) 申晓红1)† 李亚安2) 张奎3)

(西北工业大学航海学院,西安 710072)
 (西北工业大学,海洋声学信息感知工业和信息化部重点实验室,西安 710072)

3) (西安精密机械研究所, 西安 710075)

(2024年10月28日收到; 2024年12月23日收到修改稿)

复杂系统的非线性动力学研究一直是物理学领域的重点,因为其有助于提高系统建模的精度,从而实现 更准确的系统分析与预测.非线性特征参数,如熵和李雅普诺夫指数,能够有效揭示复杂系统中隐含的动力 学特性.与传统的一维时间序列数据相比,多通道阵列数据包含更多关于非线性系统动力学的信息,相空间 重构后的数据矩阵在结构上也与多通道阵列数据展现出相似的特性.然而,现有的非线性特征参数计算方法 仅适用于一维时间序列.仅选取多通道数据中的一个通道来计算非线性特征参数,会导致大量系统信息的浪 费,从而降低了非线性特征参数的性能与精度.本文针对含有相空间重构的多尺度样本熵与多尺度排列熵两 种经典非线性特征参数,利用阵列多通道数据代替算法中的相空间重构以提升算法性能.通过分析相空间重 构参数与实际阵列结构参数之间的关系,给出嵌入维数、时间延迟与阵元个数、阵元间距之间的关系.通过 构造多组仿真阵列数据,分别计算其多尺度样本熵和多尺度排列熵.结果表明,使用阵列数据代替相空间重 构可以有效地提升两种熵算法的性能.其中,使用阵列数据的多尺度样本熵算法可以更准确地揭示信号在不同时 间尺度上的复杂度.进一步地,实测阵列数据分析也与仿真结果一致,证明了阵列数据含有更丰富的系统信 息.而其隐含的更多信息可以有效地提升非线性特征参数的性能,实现对不同类型系统的准确区分.

关键词:非线性动力学,阵列数据分析,多尺度样本熵,多尺度排列熵
 PACS: 05.45.-a, 89.70.Cf, 43.60.Cg
 CSTR: 32037.14.aps.74.20241512

#### 1 引 言

非线性动力学作为一门研究非线性系统行为 的学科,近年来在科学与工程技术领域获得了广泛 的应用<sup>[1-4]</sup>.相空间重构是非线性动力学研究的一 个重要内容<sup>[5]</sup>.它为使用有限的一维时间序列数据 分析复杂系统提供了基础.对于一个非线性动力学 系统,随时间变化的系统状态是受整个系统的运动 规律支配的.由观测得到的一维时间序列隐含着整 个系统的运动规律<sup>[6-8]</sup>. 相空间重构就是利用这些 一维时间序列重构非线性系统的动力学特性, 其依 据是延迟嵌入维定理<sup>[9]</sup>. 通过选择适当的时间延迟 和嵌入维数, 将观测到的一维时间序列映射到高维 相空间中<sup>[10,11]</sup>. 相空间重构不仅可以帮助理解系统 的演化轨迹和吸引子结构, 还可以为后续的动力学 分析和系统特征参数计算提供基础. 传统的非线性 特征算法由于观测得到的一维时间序列数据不足 以完整描述系统的动力学特性, 因此需要借助相空 间重构来弥补一维时间序列数据对系统描述不充

<sup>\*</sup> 国家自然科学基金 (批准号: 11874302) 和国家自然科学基金重点项目 (批准号: 62031021) 资助的课题.

<sup>†</sup> 通信作者. E-mail: xhshen@nwpu.edu.cn

<sup>© 2025</sup> 中国物理学会 Chinese Physical Society

分的问题,从而对系统的动力学特性进行有效的表征<sup>[12-14]</sup>.通过对这些非线性特征参数进行计算和分析,进而更全面地理解系统的动力学行为,揭示系统内在的规律性和演化机制.

阵列技术在目标信号检测、方位估计等方面具 有广泛应用<sup>[15,16]</sup>, 阵列技术的发展为非线性系统研 究提供了新的思路. 利用多通道阵列数据可以对非 线性系统进行更加准确的表征, 从而获得更加丰富 的动力学信息. 然而, 在非线性动力学领域, 阵列 数据对非线性特征的影响研究还有待深入. 从理论 上分析, 虽然重构后的数据矩阵在结构上与阵列多 通道数据相似, 但是相空间重构参数与阵列多通道 数据之间的关系尚不明确. 多通道数据携带的更多 信息是否可以给非线性特征参数的计算带来性能 提升也值得研究.

本文选取多尺度样本熵<sup>[17]</sup> 与多尺度排列熵<sup>[18]</sup> 研究阵列数据对基于相空间重构的非线性算法的 影响. 作为两种经典的非线性特征算法, 多尺度样 本熵与多尺度排列熵分别从不同的角度对一维时 间序列在时域上的结构特性进行了有效的表征. 随着研究的不断深入, 有学者对这两种算法进行了改 进<sup>[19,20]</sup>, 改进后的算法被广泛应用于水声信号处 理<sup>[21]</sup>、机械故障诊断等领域<sup>[22-24]</sup>. 虽然改进后的算 法对信号结构特性的表征更加准确, 在对含噪信号 的结构特性进行表征时具有更好的鲁棒性. 但是, 改进算法中复杂的预处理步骤会对熵值曲线产生 干扰. 进而难以准确分析出阵列数据代替相空间重 构对算法产生的影响.

本文首先从多通道数据结构的角度对相空间 重构后的重构矩阵以及阵列数据进行了分析.对比 两种数据的数据结构,给出了相空间重构中参数与 阵列结构参数之间的对应关系.具体来说,对于阵 列多通道数据,阵元数与相空间重构中的嵌入维数 是等价的.而阵元间距、声速以及目标与阵列之间 的夹角会对时间延迟产生影响.进一步,使用不同 的仿真阵列数据与实际阵列数据,分别使用其中一 个通道的数据计算多尺度样本熵与多尺度排列熵 结果.再使用阵列数据直接代替相空间重构步骤计 算对应熵值.仿真与实际数据的分析结果表明,使 用阵列数据计算得到的多尺度排列熵曲线可以更 计算得到的多尺度排列熵曲线可以更清晰地揭示 信号在不同时间尺度上的复杂度结构特征,证明了 利用阵列数据代替相空间重构可以显著地提升非 线性特征的性能.本文的主要贡献如下:1)给出了 相空间重构与阵列数据结构之间的关系;2)利 用仿真实验分析证明了相空间重构后的数据结构 与阵列数据结构之间的等效性;3)给出了一种新 的基于阵列数据的非线性特征计算方法.新计算方 法不仅可以有效地提升不同类型信号的区分度,还 可以减少信号随尺度变化的波动,更清晰地揭示信 号随尺度变化的复杂度结构特征.

#### 2 多尺度样本熵与多尺度排列熵

多尺度样本熵和多尺度排列熵是在样本熵和 排列熵算法的基础上增加了粗粒化作为数据预处 理步骤<sup>[25,26]</sup>. 在样本熵算法中,首先统计了时间序 列中相似的两个长度为 m的子序列,再增加一个 样本点后仍然保持相似的条件概率,并计算了该条 件概率的负自然对数作为时间序列复杂度的估计. 而排列熵算法则将数据重构为若干个长度为 m的 子序列,并将这些子序列映射到 m维序数模式空 间中. 在序数模式空间中分析时间序列的结构特 性,以估计时间序列的复杂度.

#### 2.1 多尺度样本熵

给定一段长度为 N的一维时间序列  $\{x_i\}$ ,  $1 \leq i \leq N$ ,利用 (1) 式将  $\{x_i\}$ 分解到不同的时间 尺度上:

$$\{y_i^{(n)}\} = \frac{1}{n} \sum_{k=1}^n x_{n(i-1)+k}, \ 1 \leqslant i \leqslant N_n.$$
 (1)

式中,  $\{y_i^{(n)}\}$ 为粗粒化后的数据, 右上角标 n 为尺度因子,  $N_n$  为N/n 向下取整. 下文中使用  $\{y_i\}$ 表示粗粒化后的数据, 省略尺度因子角标 n.

得到粗粒化数据 {*y<sub>i</sub>*} 后, 给定嵌入维数 *m* 与 时间延迟 *τ*, 利用 (2) 式对粗粒化后的数据进行相 空间重构:

$$\mathbf{Y}_{m}(i) = [y_{i}, y_{i+\tau}, \cdots, y_{i+(m-1)\tau}],$$
  
$$1 \leq i \leq N_{n} - (m-1)\tau.$$
(2)

对数据进行相空间重构得到 Y<sub>m</sub>(i) 后, 定义相 空间中任意两个不同向量之间的距离:

$$d[\mathbf{Y}_{m}(i), \mathbf{Y}_{m}(j)] = \max[|y_{i+k\tau} - y_{j+k\tau}|],$$
  

$$0 \leq k \leq m-1, \ 1 \leq i, \ j \leq N_{n} - (m-1)\tau, \ i \neq j.$$
(3)

定义相空间中任意两向量之间的距离后,设定 SD×r作为阈值,其中 r是阈值系数,SD是原始 时间序列 { $x_i$ }的标准差.针对相空间中某一特定 向量  $Y_m(i)$ ,计算其与相空间中其余向量之间的距 离.统计相空间中与向量  $Y_m(i)$ 的距离小于给定阈 值 SD×r的向量个数,记为  $C_i^m(r)$ .利用 (4)式 即可计算出相空间中与向量  $Y_m(i)$ 距离小于给定 阈值的频率,记为  $B_i^m(r)$ :

$$B_i^m(r) = \frac{1}{N_n - (m-1)\tau - 1} \cdot C_i^m(r).$$
(4)

在得到针对特定向量  $Y_m(i)$  的频率  $B_i^m(r)$  后, 遍历相空间中的所有向量计算得到对应的  $B_i^m(r)$ . 对所有的  $B_i^m(r)$  求和取平均, 具体计算公式如下:

$$B^{m}(r) = \frac{1}{N_{n} - (m-1)\tau} \cdot \sum_{i=1}^{N_{n} - (m-1)\tau} B_{i}^{m}(r).$$
 (5)

得到 *B<sup>m</sup>*(*r*) 后,将嵌入维数由 *m* 变为 *m*+1. 重复 (2) 式—(5) 式计算得到对应的 *B<sup>m+1</sup>*(*r*).则时间序 列 {*x<sub>i</sub>*} 在尺度 *n*上的样本熵值为

 $MSE(m, \tau, r, n, N) = -\ln[B^{m+1}(r)/B^m(r)].$  (6)

多尺度样本熵算法的计算步骤中,分别重构 了 m 维和 m+1 维相空间.在重构出相空间后,统 计相空间中不同向量相似的频率.最终计算的多尺 度样本熵值本质上是两段长度为 m 的子序列在增 加一个样本点后,依然保持相似的条件概率的负自 然对数.由此可以看出,多尺度样本熵表征的是时 间序列在时域结构上的相关程度.虽然这种时域结 构上的相关程度可以在一定程度上侧面反映时间 序列的复杂度,但是并不能将两者混淆.

#### 2.2 多尺度排列熵

同样,考虑长度为 N的时间序列 { $x_i$ }, $1 \le i \le N$ ,利用 (1)式对数据进行粗粒化处理后得到在尺度 n上的粗粒化数据 { $y_i$ }.与多尺度样本熵算法 类似,给定嵌入维数 m与时间延迟  $\tau$ 后,利用 (2)式进行相空间重构.将相空间中所有向量中的 各个元素按升序进行排列,以向量  $Y_m(i)$ 为例,如 (7)式所示:

$$\boldsymbol{Y}_{m}(i) = [y_{i}, y_{i+\tau}, \cdots, y_{i+(m-1)\tau}] \xrightarrow{\text{ $\underline{k} \oplus \mathbb{P}^{\#} \overline{M}$}} [y_{k_{1}}, y_{k_{2}}, \cdots, y_{k_{m}}].$$

$$(7)$$

式中,  $k_1, k_2, \dots, k_m$ 分别表示各元素的原始位置. 通过建立相空间中向量与序数模式 $\pi_i = [k_1, m_i]$  *k*<sub>2</sub>,...,*k<sub>m</sub>*]的对应关系,将相空间中的所有向量映 射到序数模式空间中.一个 *m* 维的序数模式空间 中共有 *m*!种可能的序数模式.例如,当*m* = 3时, 6 种不同的序数模式的示意图如图 1 所示.



图 1 m = 3时6种不同的序数模式示意图,黑色圆点代 表数据点,虚线连接的数据点之间的高低关系代表数据点 之间的大小关系

Fig. 1. All possible ordinal patterns when m=3, the black dots represent data points, and the dashed lines connecting the data points indicate the relative magnitude between them.

将相空间中所有向量映射到序数模式空间中 后,统计每种序数模式出现的频次 h<sub>l</sub>,并计算其出现 的概率 p<sub>l</sub>=h<sub>l</sub>/[N<sub>n</sub>-(m-1)τ],其中 l=1,2,...,m!. 则多尺度排列熵的定义如下:

$$MPE(m, \tau, n, N) = -\sum_{l=1}^{m!} p_l \ln(p_l).$$
(8)

多尺度排列熵对不同时间尺度上时间序列的 复杂程度进行了表征. 当相空间中仅存在一种序数 模式时 (时间序列单调递增或单调递减), 多尺度 排列熵取最小值 0; 当序数模式服从均匀分布时,  $p_l = 1/m!, l = 1, 2, \dots, m!, 多尺度排列熵取最大$ 值 ln(<math>m!). 因此在计算过程中, 可以对多尺度排列 熵值进行归一化处理. 后文如无特殊说明, 多尺度 排列熵值均为归一化后的结果.

#### 3 数据结构分析

由 (2) 式可知, 多尺度样本熵与多尺度排列熵 算法在计算过程中将一维的时间序列重构为了一 个 $m \times [N - (m - 1)\tau]$ 的矩阵 (尺度为1时).将 (2) 式等号右侧展开如下:

$$\begin{bmatrix} y_1, y_{1+\tau}, \cdots, y_{1+(m-1)\tau} \\ [y_2, y_{2+\tau}, \cdots, y_{2+(m-1)\tau}] \\ \cdots \\ [y_{N_n - (m-1)\tau}, y_{N_n - (m-2)\tau}, \cdots, y_{N_n}] \end{bmatrix}.$$
 (9)

(9) 式中的矩阵,每一行是一个相空间中维数为 m的向量,而每一列则是时间序列 {y<sub>i</sub>} 的一段子 序列.并且在同一行中,相邻两列的采样点相差固 定的时间延迟 τ.这与由阵列采集到的多通道数据 呈现出相似的结构特性.以水下方位估计中的均匀 线列阵为例,假设阵元个数为 m<sub>a</sub>,阵间距为 d. 在 远场条件下目标与阵列中心垂线的夹角为 θ,水下 声速假设为 c,采样率为 f<sub>s</sub>.线列阵结构示意图如 图 2 所示.



图 2 远场条件下的均匀线列阵结构示意图 Fig. 2. Schematic diagram of a uniform linear array structure under far-field conditions.

由图 2 可以看出, 在远场条件下各个阵元的垂 线方向与目标来波的方向夹角均为 θ. 在阵间距、 采样率和声速均已知的情况下, 可以根据 (10) 式 求出相邻阵元之间的采样间隔 τ<sub>a</sub>:

$$\tau_{\rm a} = \frac{d \cdot \sin \theta}{c} \cdot f_{\rm s}.$$
 (10)

结合 (9) 式和图 2, 可以观察到相空间重构后的数 据结构与由均匀线列阵采集的数据结构相似. 需要 特别注意的是, 与多尺度排列熵算法不同, 多尺度 样本熵算法中包含了增加嵌入维数的步骤. 具体来 说, 对于多尺度排列熵算法, 相空间重构中的参数 m等效于均匀线列阵的阵元数  $m_a$ . 而对于多尺度 样本熵算法, 初始相空间重构中的 m 与阵元个数  $m_a$ 之间的关系是  $m = m_a - 1$ . 对于这两种熵算法 来, 参数  $\tau$  与阵间距 d 可以通过 (10) 式相互转化. 这意味着, 在多尺度样本熵和多尺度排列熵算法 中, 相空间重构这一步骤本质上将一段一维的时间 序列扩展为一个 m 维的阵列数据.

4 仿真实验分析

为了对比研究单通道数据与阵列数据对非线 性特征参数的影响,构造了由均匀线列阵采集到的 远场多通道数据.首先,直接使用其中一个通道的 数据,并通过相空间重构技术将一维时间序列重构 为多维相空间矩阵,再计算其多尺度样本熵与多尺 度排列熵.接着,跳过多尺度样本熵与多尺度排列 熵算法中的相空间重构步骤.直接使用阵列数据代替多维相空间矩阵,再计算多尺度样本熵与多尺度 排列熵值.

参考图 2, 假设均匀线列阵阵元个数为 5, 阵 间距为 5 m, 水中声速假设为 1500 m/s, 目标信号 假定为含有 80 Hz, 210 Hz, 430 Hz 线谱成分的连 续波信号. 水听器采样率为 5000 Hz, 目标与阵列 垂直方向夹角设置为 30°. 根据上述参数与 (10) 式 的分析, 可以得出以下结论: 在使用多尺度样本熵 时, 该均匀线列阵采集到的多通道数据与嵌入维数 m = 4, 时间延迟  $\tau \approx 8$  的相空间重构数据具有相 似的结构特征. 而在使用多尺度排列熵时, 该阵列 数据与嵌入维数 m = 5, 时间延迟  $\tau \approx 8$  的相空间 重构数据具有相似的结构特征. 为了模拟实际环 境, 在目标信号中加入高斯白噪声使信噪比为 0 dB. 分别使用多尺度样本熵与多尺度排列熵算法计算 熵值, 并对比计算高斯白噪声的熵值, 以进行进一 步的研究.

#### 4.1 基于单通道数据的熵特征提取

原始的多尺度样本熵和多尺度排列熵算法在 计算过程中只需要使用单通道数据.因此,在第一 组仿真实验中,仅使用一个阵元采集到的一维时间 序列进行分析.相空间重构的参数设置遵循与实际 阵列等效的参数设置,即 $m_{MSE} = 4$ ,  $m_{MPE} = 5$ ,  $\tau = 8, r = 0.15$ .其中 $m_{MSE}$ 指多尺度样本熵算法中 的嵌入维数,而 $m_{MPE}$ 指多尺度排列熵的嵌入维数. 数据长度为10000,尺度因子范围为[1,15].对含噪 信号和高斯白噪声分别独立重复进行10次试验, 得到多尺度样本熵与多尺度排列熵的误差图结果. 具体如图 3 所示.

从图 3(a) 可以观察到, 在使用单通道数据计 算多尺度样本熵时, 由于信噪比较低, 含噪目标信 号的熵值曲线与高斯白噪声的熵值曲线相似. 随着 尺度因子的增大, 两种曲线逐渐减小. 值得注意的 是, 含噪信号的熵值结果方差明显大于高斯白噪声 的熵值结果方差. 因此, 在计算多尺度样本熵时, 仅使用单通道数据难以有效区分两种不同的信号. 然而, 从图 3(b) 的多尺度排列熵结果可以看出, 随 着尺度因子的增大, 含噪目标信号的熵值出现明显 的波动. 整体上, 含噪目标信号的熵值小于高斯白 噪声的熵值. 因此, 多尺度排列熵能够有效地区分 这两种信号.





#### 4.2 基于阵列数据的熵特征提取

在使用阵列数据计算多尺度样本熵和多尺度 排列熵的结果时,相较于直接使用单通道数据计 算,计算过程变得更为复杂.从上述分析可以看出, 利用阵列数据可以直接替代原始多尺度样本熵和 多尺度排列熵算法中的相空间重构步骤.直观上, 这似乎只是一个简单的数据替代问题.然而,在仿 真实验中,还有以下几个问题需要额外关注.

首先需要考虑的是加入高斯白噪声时遇到的问题.在构造阵列仿真数据时,我们要在目标信号中加入一定信噪比的高斯白噪声.当仅使用一个通道的数据来计算多尺度样本熵和多尺度排列熵时,可以直接将高斯白噪声加入生成的目标信号中.然而,在阵列数据的情况下,构造仿真阵列数据的过程是先生成一段信号,然后利用滑动窗口每隔时间间隔 τ 截取出每个通道的数据.因此,确定何时加入高斯白噪声变得尤为关键.

当使用相空间重构技术将一维数据扩展为阵 列数据时,由于一维数据事先已经加入了高斯白噪 声,因此在重构后的数据中相邻两列的采样点应严 格满足关系  $y_{i+\tau,j} = y_{i,j+1}$ .其中, $y_{i,j}$ 表示相空间 重构后矩阵中第 i行第 j列的样本点(参考(9)式). 然而,在实际阵列数据采集过程中,每个通道接收 到的高斯白噪声是相对独立的,因此不满足上述关 系.为了更好地模拟实际采集到的阵列数据,在构 造阵列仿真数据时首先构造具有固定时间延迟的 阵列数据,然后为每一个通道分别加入相同功率的 高斯白噪声.

2) 接下来需要考虑的是粗粒化过程.在使用
 单通道数据计算多尺度样本熵和多尺度排列熵时,

首先对数据进行粗粒化处理,将数据分解到不同的 时间尺度.然后,在这些不同的时间尺度上进行相 空间重构,并计算相应的熵值.然而,在实际的阵 列数据采集过程中,首先获取整个阵列的数据,接 着需要对每一个通道的数据进行粗粒化预处理.值 得注意的是,对于每个通道的数据,在进行粗粒化 处理之前,需要先加入相应通道的高斯白噪声.

为了更好地对比使用单通道数据和仿真阵列 多通道数据计算的熵结果,设计了两组多通道数据 的仿真实验. 第一组实验中的阵列多通道数据是按 照相空间重构的逻辑严格构造的.首先,生成一个 包含 80 Hz, 210 Hz, 430 Hz 线谱成分的连续波信 号,连续波信号的长度需大于10000.在该信号中 加入信噪比为0dB的高斯白噪声.接着,按照不 同的尺度因子 n 对一维数据进行粗粒化处理, 并使 用长度为 |10000/n | 的滑动窗将一维数据转换为阵 列多通道数据, 1-1 表示向下取整. 使用这些构造的 阵列多通道数据替代多尺度样本熵算法和多尺度 排列熵算法中的相空间重构步骤,并计算相应的熵 值. 其余未额外说明的计算参数与图 3 的仿真实验 设置相同,实验结果如图4所示.将图4的结果与 图 3 进行对比,可以看出无论是多尺度样本熵还是 多尺度排列熵,其熵值曲线的整体趋势和波动都与 单通道情况下的熵值曲线相一致. 通过对比图 4 与 图 3 的结果, 证明了相空间重构技术本质上是将一 维的时间序列重构为多维阵列数据,从而弥补数据 获取量不足的缺陷,进而能够对复杂系统进行有效 的非线性动力学分析.

第二组多通道数据的仿真实验在构造仿真阵 列数据时,采取了更符合实际情况的构造方法.首

先生成含有 80 Hz, 210 Hz, 430 Hz 线谱成分的一 维连续波信号,信号长度大于10000.利用长度为 10000的滑动窗将一维数据构造为阵列多通道数 据,给每个通道加入信噪比为0dB的高斯白噪声. 得到含噪的阵列连续波信号数据后,利用粗粒化过 程处理每一个通道的数据. 最后, 直接利用粗粒化 后的多通道阵列数据计算对应的多尺度样本熵与 多尺度排列熵结果.其余计算参数均保持不变,具 体结果如图 5 所示.图 5(a) 中的多尺度样本熵曲 线与图 4(a) 中的多尺度样本熵曲线相比, 虽然含 噪信号的方差增大,但是含噪信号和高斯白噪声熵 值曲线的均值差更大,且熵值随尺度变化的趋势更 加平缓,从而更好地实现了对不同类型信号的区 分. 而图 5(b) 中的多尺度排列熵曲线与图 4(b) 中 的多尺度排列熵曲线相比,含噪信号的熵值曲线有 较大差异. 虽然图 4(b) 和图 5(b) 中的含噪信号熵 值曲线整体都呈现出下降的趋势,但是图 5(b)中 的曲线波动幅度更大. 且图 5(b) 中含噪信号的熵

值曲线在尺度 10 出现局部极小值, 而在图 4(b) 中 在尺度 6, 9, 12 处出现局部极小值. 两幅图并没有 明显的对应关系.

多尺度排列熵算法与多尺度样本熵算法相比, 对数据长度的要求更低.因此在计算多尺度排列熵 时,通常可以选取较大的尺度范围从而更全面地分 析数据在不同时间尺度上的复杂度结构特性.因 此,本文分别在图 4(b) 和图 5(b) 的仿真实验基础 上改变计算参数.将每个通道的数据长度增大到 50000,尺度范围从[1,15]增大到[1,50],其余参数 不变.具体多尺度排列熵结果如图 6 所示.

对比图 6 的两个多尺度排列熵结果可以看出, 当尺度范围增大后,信号在不同时间尺度上的结构 特征被更完整的揭示出来.图 6(a),(b)的含噪信 号多尺度排列熵结果曲线都呈现出先减小后增大 的趋势.在图 6(b)中,含噪信号的曲线在尺度[1,9] 和 [11,24]范围内逐渐减小,在尺度[26,44]和[46,50] 范围内逐渐增大.并且,曲线分别在尺度 10,25,45



图 4 使用按照相空间重构逻辑构造的阵列数据的多尺度样本熵与多尺度排列熵算法结果 (a) 多尺度样本熵; (b) 多尺度排列熵 Fig. 4. The multiscale sample entropy and multiscale permutation entropy results using array data constructed according to phase space reconstruction technique: (a) Multiscale sample entropy result; (b) multiscale permutation entropy result.



图 5 仿真阵列数据的多尺度样本熵与多尺度排列熵算法结果 (a)多尺度样本熵; (b)多尺度排列熵

Fig. 5. The multiscale sample entropy and multiscale permutation entropy results for simulated array data: (a) Multiscale sample entropy result; (b) multiscale permutation entropy result.

处出现局部极小值. 反观基于相空间重构的多尺度 排列熵曲线结果, 虽然整体上曲线的趋势也呈现出 先减小后增大的特征, 并且在尺度 25 处出现局部 极小值. 但是由于曲线整体不稳定导致无法有效的 对数据在不同时间尺度上的结构特征进行分析, 并 且无法清晰地观察到位于尺度 10 和 45 处的局部 极小值.

5 实际阵列数据分析

本文使用 3 种由线列阵采集到的不同类型的 实际多通道数据, 分别为两类不同的舰船辐射噪声 和海洋背景噪声. 线列阵的阵元个数为 5, 阵间距 为 5 m, 采样频率为 9960 Hz. 由于实测数据无法 精确获得目标与阵列垂直方向的夹角, 进而无法计 算出理论上相邻阵元之间的时间延迟, 因此采用互 相关函数计算两类舰船辐射噪声多通道数据中相 邻阵元的时间延迟. 通过计算得到相邻通道间数据 的互相关函数结果. 互相关函数取最大值时对应的 时间延迟即认为是相邻阵元间的时间延迟. 表 1 列 出了计算结果.

表 1 舰船辐射噪声相邻阵元的时间延迟 Table 1. Time delay of adjacent array elements of ship-radiated noise.

舰船 类型	通道1- 通道2	通道2- 通道3	通道3- 通道4	通道4- 通道5	平均时间 延迟
舰船1	22	23	22	18	21.25
舰船2	20	21	21	17	19.75

海浪波动以及人为测量误差等因素导致相邻 通道之间的时间延迟不严格相等,因此计算所有通

1.00

0.95

0.90

0.85

0.75

0.70

0.65

多尺度排列熵值

道间时间延迟的均值并取整.综上,在计算多尺度 样本熵时,舰船1的阵列数据近似等效为嵌入维数 m = 4,时间延迟 $\tau \approx 21$ 的相空间重构后数据;舰 船2的阵列数据近似等效为嵌入维m = 4,时间延 迟 $\tau \approx 20$ 的相空间重构后数据.而在计算多尺度排 列熵时,舰船1的阵列数据近似等效为嵌入维数 m = 5,时间延迟 $\tau \approx 21$ 的相空间重构后数据;舰 船2的阵列数据近似等效为嵌入维m = 5,时间延迟 $\tau \approx 20$ 的相空间重构后数据.

首先, 仅利用其中一个通道的数据直接计算多 尺度样本熵与多尺度排列熵结果, 数据长度为 20000, 多尺度样本熵的尺度因子范围为 [1,15], 多 尺度排列熵的尺度因子范围为 [1,50]. 相空间重构 的参数设置遵循与实际阵列等效的参数设置, 即 $m_{MSE} = 4$ ,  $m_{MPE} = 5$ ,  $\tau_1 = 21$ ,  $\tau_2 = 20$ ,  $\tau_n = 1$ , r = 0.15. 其中,  $\tau_1$  是舰船 1 的时间延迟,  $\tau_2$  是舰 船 2 的时间延迟, 而 $\tau_n$ 是环境噪声的时间延迟. 重 复 10 次实验, 绘制多尺度样本熵与多尺度排列熵 的误差图, 结果如图 7 所示.

图 7(a) 的多尺度样本熵结果表明,多尺度样 本熵可以有效区分舰船与环境噪声. 然而, 仅依靠 多尺度样本熵无法有效区分不同类型的舰船辐射 噪声. 虽然当尺度大于 8 时, 两类舰船辐射噪声的 多尺度样本熵曲线开始出现分离. 但由于两类舰船 辐射噪声的多尺度样本熵曲线的整体均值过于相 似, 在实际应用中, 仅使用一个通道的数据计算得 到的多尺度样本熵曲线无法有效的对不同类型的 舰船辐射噪声进行区分. 对于多尺度排列熵, 仅使 用一个通道数据计算得到的多尺度排列熵曲线也



图 6 多尺度排列熵算法结果 (a) 基于相空间重构构造的阵列数据的多尺度排列熵算法结果; (b) 仿真阵列数据的多尺度排列 熵算法结果

Fig. 6. Multiscale permutation entropy results: (a) Array data constructed according to phase space reconstruction technique; (b) simulated array data.

同样并不理想.信号的剧烈波动使得无法通过多尺 度排列熵曲线区分两类不同的舰船.虽然在低尺度 下,环境噪声与舰船辐射噪声的区别相对明显,但 随着尺度的增大,无法有效区分舰船与环境噪声.

接下来,使用由阵列采集到的多通道数据.数 据为5通道数据,每个通道内有20000样本点.采 用与图5仿真实验相同的计算思路,利用多通道数 据代替多尺度样本熵与多尺度排列熵算法中的相 空间重构步骤,计算不同类型实际多通道数据的多 尺度样本熵值与多尺度排列熵值.多尺度样本熵尺 度范围为[1,15],多尺度排列熵的尺度因子范围为 [1,50],阈值系数r = 0.15.重复进行10次实验,结 果如图8所示.

对比图 8(a) 与图 7(a),可以看到使用阵列数 据计算得到的两类舰船辐射噪声的多尺度样本熵 结果在尺度大于 1 时就有明显的区别.在图 8(a) 的虚线矩形框中,两种舰船辐射噪声的熵值曲线有 明显区别.而图 7(a)中,在尺度小于 8 时利用多尺 度样本熵曲线是无法区分两种舰船辐射噪声的.并 且可以注意到,图 8(a)使用阵列数据计算得到的 熵值曲线在均值、方差以及整体随尺度变化的趋势 并没有出现明显变化的基础上对不同类型信号的 区分度更好. 实测数据分析结果与仿真实验中的结 果 (图 3(a), 图 4(a), 图 5(a)) 相一致. 这说明, 利 用阵列数据代替相空间重构计算多尺度样本熵时, 阵列数据中更丰富的信息使得算法可以有效的对 两种相似的信号进行区分. 而对比图 8(b) 与图 7(b) 的多尺度排列熵曲线可以看出,利用图 8(b) 的多 尺度排列熵曲线可以有效区分舰船1,舰船2与环 境噪声. 图 7(b) 中两类舰船辐射噪声的熵值曲线 近乎完全重叠,并且随尺度变化出现了剧烈波动. 图 8(b) 中的熵值曲线明显比图 7(b) 的更加平稳. 在图 8(b) 的虚线框中, 三类信号的熵值曲线也有 明显的差异. 这说明使用阵列数据计算得到的多尺 度排列熵曲线对三类实测数据的区分度更高. 当尺 度较大时,虽然数据长度缩短,但是使用阵列数据



图 7 使用一个通道数据计算实际数据的多尺度样本熵与多尺度排列熵结果 (a) 多尺度样本熵结果; (b) 多尺度排列熵结果 Fig. 7. The multiscale sample entropy and multiscale permutation entropy results of the actual data calculated using only one channel data: (a) Multiscale sample entropy results; (b) multiscale permutation entropy results.



图 8 使用阵列数据代替相空间重构计算的多尺度样本熵与多尺度排列熵结果 (a) 多尺度样本熵结果; (b) 多尺度排列熵结果 Fig. 8. The results of multiscale sample entropy and multiscale permutation entropy calculated by array data: (a) Multiscale sample entropy results; (b) multiscale permutation entropy results.

计算得到的多尺度排列熵曲线方差较小. 这也说明 了使用阵列数据计算得到的多尺度排列熵结果更 加稳定.

#### 6 结 论

本文对比了相空间重构后的数据矩阵与阵列 采集到的多通道数据在结构上的特征,并定量分析 了相空间重构参数与实际阵列参数之间的对应关 系. 在仿真实验中, 首先利用仿真阵列数据的其中 一个通道计算了多尺度样本熵和多尺度排列熵曲 线.结果显示,当信噪比为0dB时,多尺度样本熵 仅用一个通道数据难以有效区分含噪目标信号和 高斯白噪声,而多尺度排列熵由于与噪声曲线均值 存在显著差异, 仅用一个通道数据即可有效区分两 种不同的信号. 接下来, 构造了多通道仿真数据, 确保其与通过相空间重构生成的单通道数据在结 构上保持一致. 将这些多通道数据直接代替了多尺 度样本熵和多尺度排列熵算法中的相空间重构,并 计算了相应的熵值曲线. 结果证实了在特定条件 下,相空间重构技术与阵列采集的多通道数据是等 价的. 进一步地, 构造了更符合实际情况的阵列数 据,并分别使用多尺度样本熵和多尺度排列熵算法 计算了含噪目标信号与高斯白噪声的熵值.对比分 析表明,由于阵列数据包含更多的目标系统信息, 使用阵列数据代替相空间重构步骤可以有效提升 算法性能.

本文还使用了3组不同类型的多通道实际数据,分别利用多尺度样本熵与多尺度排列熵进行分析.结果表明,仅使用单个通道数据计算得到的多 尺度样本熵与多尺度排列熵曲线无法有效的对不 同类型的实际数据进行区分.而直接使用阵列数据 代替相空间重构后,多尺度样本熵和多尺度排列熵 曲线均可以有效的区分不同类型的实际数据.证明 了多通道数据含有的丰富信息可以有效地提升传 统基于单通道数据与相空间重构技术的非线性特 征算法性能.

#### 参考文献

- Chen H T, Li L L, Shang C, Huang B 2022 IEEE T. Cybernetics 53 4259
- [2] Richard B, Kerry A E, Zhang F 2019 Science 363 342
- [3] Hsieh D A 1991 J. Financ. 46 1839
- [4] Xu F, Zou Z J, Yin J C, Cao J 2013 Ocean Eng. 67 68
- [5] Takens F 1980 Dynamical Systems and Turbulence (Heidelberg: Springer Berlin) p366
- [6] Liu B Z, Peng J H 2004 Nonlinear Dynamics pp389-403 (in Chinese) [刘秉正, 彭建华 2004 非线性动力学 (北京: 高等教育 出版社) 第 389-403 页]
- [7] Zhao J Y, Jin N D 2012 Acta Phys. Sin. 61 094701 [赵俊英, 金宁德 2012 物理学报 61 094701]
- [8] Huang Z H, Li Y A, Chen Z, Liu L 2020 Acta Phys. Sin. 69 160501(in Chinese) [黃泽徽, 李亚安, 陈哲, 刘恋 2020 物理学 报 69 160501]
- [9] Grigoryeva L, Hart A, Ortega J P 2021 Phys. Rev. E 103 062204
- [10] Kennel M B, Brown R, Abarbanel H 1992 Phys. Rev. A 45 3403
- [11] Zhang S Q, Jia J, Gao M, Han X 2010 Acta Phys. Sin 59
   1576(in Chinese) [张淑清, 贾健, 高敏, 韩叙 2010 物理学报 59
   1576]
- [12] Richman J S, Moorman J R 2000 Am. J. Physiol-Heart. C 278 H2039
- [13] Bandt C, Pompe B 2002 Phys. Rev. Lett 88 174102
- Bryant P, Brown R, Abarbanel H 1990 Phys. Rev. Lett. 65 1523
- [15] Ferrara E, Parks T 1983 IEEE T. Antenn. Propag. 31 231
- [16] He J, Liu Z 2009 Signal Process. 89 1715
- [17] Costa M, Goldberger A L, Peng C K 2005 Phys. Rev. E 71 021906
- [18] Liu T B, Yao W P, Wu M, Shi Z R, Wang J, Ning X B 2017 *Physica A* 471 492
- [19] Humeau-Heurtier A, Wu C W, Wu S D 2015 IEEE Signal Proc. Let. 22 2364
- [20] Amigó J M, Dale R, Tempesta P 2021 Chaos 31 013115
- [21] Li W J, Shen X H, Li Y A 2019 Entropy-Switz 21 793
- [22] Li Y B, Xu M Q, Wang R X, Huang W H 2016 J. Sound Vib. 360 277
- [23] Gao S Z, Wang Q, Zhang Y M 2021 IEEE T. Instrum. Meas. 70 3514908
- [24] He L, Du L, Zhuang Y Q, Li W H, Chen J P 2008 Acta Phys. Sin. 57 6545 (in Chinese) [何亮, 杜磊, 庄奕琪, 李伟华, 陈建平 2008 物理学报 57 6545]
- [25]~ Zhang Y 1991 J. Phys. I 1 971
- [26] Fogedby H C 1992 J. Stat. Phys. 69 411

### Extraction of nonlinear feature parameters based on multi-channel dataset<sup>\*</sup>

LI Weijia<sup>1)2)</sup> SHEN Xiaohong<sup>1)†</sup> LI Yaan<sup>2)</sup> ZHANG Kui<sup>3)</sup>

1) (School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China)

2) (Key Laboratory of Ocean Acoustics and Sensing, Ministry of Industry and Information Technology,

Northwestern Polytechnical University, Xi'an 710072, China)

3) (Xi'an Precision Machinery Research Institute, Xi'an 710075, China)

(Received 28 October 2024; revised manuscript received 23 December 2024)

#### Abstract

Phase space reconstruction plays a pivotal role in calculating features of nonlinear systems. By mapping one-dimensional time series onto a high-dimensional phase space using phase space reconstruction techniques, the dynamical characteristics of nonlinear systems can be revealed. However, existing nonlinear analysis methods are primarily based on phase space reconstruction of single-channel data and cannot directly utilize the rich information contained in multi-channel array data. The reconstructed data matrix shows the structural similarities with multi-channel array data. The relationship between phase space reconstruction and array data structure, as well as the gain in nonlinear features brought by array data, has not been sufficiently studied. In this paper, two classical nonlinear features: multiscale sample entropy and multiscale permutation entropy are adopted. The array multi-channel data are used to replace the phase space reconstruction step in algorithms so as to enhance the algorithmic performance. Initially, the relationship between phase space reconstruction parameters and actual array structures is analyzed, and conversion relationships are established. Then, multiple sets of simulated and real-world array data are used to evaluate the performances of the two entropy algorithms. The results show that substituting array data for phase space reconstruction effectively improves the performances of both entropy algorithms. Specifically, the multiscale sample entropy algorithm, when applied to array data, allows for distinguishing between noisy target signals from background noise at low signal-to-noise ratios. At the same time, the multiscale permutation entropy algorithm using array data reveals the complex structure of signals on different time scales more accurately.

Keywords: nonlinear dynamics, array data analysis, multiscale sample entropy, multiscale permutation entropy

**PACS:** 05.45.-a, 89.70.Cf, 43.60.Cg

**DOI:** 10.7498/aps.74.20241512

**CSTR:** 32037.14.aps.74.20241512

<sup>\*</sup> Project supported by the National Natural Science Foundation of China (Grant No. 11874302) and the Key Program of National Natural Science Foundation of China (Grant No. 62031021).

<sup>†</sup> Corresponding author. E-mail: xhshen@nwpu.edu.cn

# 物理学报Acta Physica Sinica





Institute of Physics, CAS

#### 基于阵列多通道数据的非线性特征参数提取

李惟嘉 申晓红 李亚安 张奎

## Extraction of nonlinear feature parameters based on multi-channel dataset

LI Weijia SHEN Xiaohong LI Yaan ZHANG Kui

引用信息 Citation: Acta Physica Sinica, 74, 040501 (2025) DOI: 10.7498/aps.74.20241512 CSTR: 32037.14.aps.74.20241512 在线阅读 View online: https://doi.org/10.7498/aps.74.20241512 当期内容 View table of contents: http://wulixb.iphy.ac.cn

#### 您可能感兴趣的其他文章

#### Articles you may be interested in

一种无偏差的多通道多尺度样本熵算法

Unbiased multivariate multiscale sample entropy

物理学报. 2024, 73(11): 110502 https://doi.org/10.7498/aps.73.20231133

心律失常的多尺度建模、计算与动力学理论进展综述

A review of advances in multiscale modelings, computations, and dynamical theories of arrhythmias 物理学报. 2024, 73(21): 218702 https://doi.org/10.7498/aps.73.20240977

光纤探针测量及多尺度熵鉴别超临界类沸腾传热模式

Measurements and identification of supercritical pseudo-boiling heat transfer modes based on fiber optic probes and multiscale entropy

物理学报. 2023, 72(5): 057801 https://doi.org/10.7498/aps.72.20222060

ENSO气温关联网络结构特征差异及成因分析 Analysis of structural differences and causes of ENSO temperature network 物理学报. 2021, 70(24): 249201 https://doi.org/10.7498/aps.70.20210825

非线性超声射频信号熵对乳腺结节良恶性的定征

Classification of benign and malignant breast masses using entropy from nonlinear ultrasound radiofrequency signal 物理学报. 2021, 70(8): 084302 https://doi.org/10.7498/aps.70.20201919

星载电子器件温控的系统多尺度分析

System multi-scale analysis of temperature control for spaceborne electronic devices 物理学报. 2024, 73(18): 184401 https://doi.org/10.7498/aps.73.20240685