第 15 卷 第 8 期　　高能物理与核物理　　Vol. 15, No. 8
1991 年 8 月　　HIGH ENERGY PHYSICS AND NUCLEAR PHYSICS　　Aug., 1991

自相似随机级联过程中的条件熵和信息量*

吴元芳　刘连寿

（华中师范大学粒子物理研究所，武昌 430070）

摘　要

本文建议用相邻两代分割的条件熵和信息量来分析多重产生中的随机级联机制。证明了当所研究事件的多重数足够高时，这些特征量对分析自相似级联机制的存在及其性质很有效。

最近几年，阶乘矩随快度间隔的减小成反常指数增加的规律[1]已经在所有高能碰撞实验的分辨本领所容许的范围内观察到[2]．这个被称为“间歇”的现象吸引了大批粒子物理学家的兴趣。

各种模型被引进来解释这一现象，其中大部分都包含有一个基本的自相似级联过程。从目前情况看，自相似随机级联是最自然、最有希望在多重产生中产生间歇现象的机制．因此，寻找直接观察和检验这种机制的方法是一个非常重要而有意义的工作。本文试图运用熵和信息量的概念来达到这个目的。

假设在所考虑的快度区间 ΔY 中，所产生的粒子数 K 足够大，以致在子分割中粒子落入某个子间隔 i 中的几率可以用这个子间隔中的粒子数 K_i 和 K 的比值来表示，即：

$$p_i \simeq \frac{K_i}{K} \quad (K \gg 1). \tag{1}$$

在实际中如何运用这个假设将在本文的最后讨论。为了简单起见，以下将所有的约等号“$\simeq$”写作等号“$=$”，但请记住所有的等式仅在 $K \to \infty$ 时才严格成立。

让我们来把快度间隔 ΔY 先分成两份：δ_1，δ_2；然后再把这两份分成四份：δ_{11}，δ_{12}，δ_{21}，δ_{22}；依次类推下去．对应子间隔中的粒子数分别为：$K;K_1,K_2;K_{11},K_{12},K_{21},K_{22},\cdots$，几率分别为：

$$\begin{gathered} p_0 = 1; \\ p_1 = K_1/K,\ p_2 = K_2/K; \\ p_{11} = K_{11}/K,\ p_{12} = K_{12}/K,\ p_{21} = K_{21}/K,\ p_{22} = K_{22}/K; \\ \cdots\cdots \end{gathered} \tag{2}$$

本文 1991 年 1 月 3 日收到．
* 国家自然科学基金资助课题．

明显地有:

$$K_1 + K_2 = K;\ K_{11} + K_{12} = K_1;\ K_{21} + K_{22} = K_2;\cdots$$

$$p_1 + p_2 = 1;\ p_{11} + p_{12} + p_{21} + p_{22} = 1;\cdots \tag{3}$$

用 $a_j(j = 1,11,21\cdots)$ 来表示粒子落入一个基元分割中某一个子间隔(例如,左边的子间隔)中的几率,则有:

$$\begin{aligned} p_1 &= a_1, & p_2 &= 1 - a_1, \\ p_{11} &= a_{11}p_1, & p_{12} &= (1 - a_{11})p_1, \\ p_{21} &= a_{21}p_2, & p_{22} &= (1 - a_{21})p_2. \\ &\cdots & &\cdots \end{aligned} \tag{4}$$

一个基元分割的熵[3]被定义为:

$$S_j \equiv S(a_j) = -(a_j \ln a_j + (1 - a_j)\ln(1 - a_j)). \tag{5}$$

它标志在这个基元分割中粒子分布的不确定性. 明显地, 如果在所考虑的基元分割中只有统计涨落,则 a_j 和 $1\text{-}a_j$ 都等于 1/2, S_j 取它的最大值:

$$S_{j\max} = \ln 2 = 0.693. \tag{6}$$

相反地,如果在基元分割中有动力学起伏,则 a_j 将偏离 1/2,使得熵 S_j 小于它的最大值 $S_{j\max}$. 特别是对于有最大涨落的情况:

$$a_j = 1,\ 1 - a_j = 0.\ \text{或}\ a_j = 0,\ 1 - a_j = 1,$$

熵取它的极小值:

$$S_{j\min} = 0. \tag{7}$$

让我们来考虑两个相邻代的分割,例如,第一代和第二代. 对应的熵是:

$$S^{(1)} = -(p_1 \ln p_1 + p_2 \ln p_2);$$

$$S^{(2)} = -(p_{11}\ln p_{11} + p_{12}\ln p_{12} + p_{21}\ln p_{21} + p_{22}\ln p_{22}). \tag{8}$$

我们所考虑的级联过程具有以下性质: 如果知道下一代某一粒子落入某个子间隔中,那么,就可以肯定它上一代在那个间隔中. 这意味着后一代的信息中包括所有前一代的信息;换言之,上一代得到的所有信息对下一代都有用. 因此,第一代分割的熵 $S^{(1)}$,即包含在这一代中的信息量,等于第二代中所包含的第一代的信息量[4] $I(1,2)$.

$$S^{(1)} = I(1,2). \tag{9}$$

$I(1,2)$ 也称为第一代和第二代之间的相互信息[3,5],在信息论中有一个著名的公式

$$I(\alpha,\beta) = S^{(\beta)} - S_\alpha(\beta),$$

其中, α 和 β 代表两个随机实验, $S(\beta)$ 是实验 β 的熵, $S_\alpha(\beta)$ 是在实验 α 的条件下,实验 β 的条件熵. $I(\alpha,\beta)$ 是实验 α 中所包含的实验 β 的信息量. 将上式和(9)式结合有:

$$S_1(2) = S^{(2)} - I(1,2) = S^{(2)} - S^{(1)}, \tag{10}$$

其中, $S_1(2)$ 是第二代相对于第一代的平均条件熵,它标志在扣除第一代的信息后下一代分割中所剩余的信息量. 将(8)式和(4)式代入(10)式得

$$S_1(2) = S^{(2)} - S^{(1)} = p_1 S(a_{11}) + p_2 S(a_{21}). \tag{11}$$

到目前为止,所进行的讨论是一般性的,并没有用到级联过程的自相似特性. 自相似级联过程表现在,所有基元分割的几率 a_j 是服从同一几率分布的随机变量. 这意味着不

同代的分割是统计独立的，后一代的几率 a_i 与前一代无关，在这种情况下，对于式(11)作平均，并且利用(3)式的归一化条件有：

$$\langle S_1(2)\rangle = \langle S^{(2)} - S^{(1)}\rangle = \langle S_i\rangle.$$

上面的讨论可以直接推广到任意相邻两代的分割，所以有：

$$\langle \Delta S\rangle = \langle S_i\rangle, \tag{12}$$

其中，$\langle \Delta S\rangle$ 是任意相邻两代分割的平均熵差。这个重要的式子告诉我们，对于一个自相似级联过程，任意相邻两代的平均熵差等于基元分割的平均熵。这样我们得到了直接观测和检验自相似级联机制的方法：在 $K \gg 1$ 保证等式(1)近似成立的条件下，把所讨论的快度区域 ΔY 分成 2, 4, … 个子区间，看粒子在每个子分割中落入子间隔中的几率 a_i；根据(5)式计算每个子分割的熵 $S(a_i)$，如果在所考虑的过程中的确存在自相似级联，则所有的 a_i 将有相同的分布。同样，所有的 $S(a_i)$ 也相同。一个简单而且直接的检验方法是计算任意相邻两代的平均熵差并考察(12)式的有效性。

以下我们采用一个简单的自相似级联模型[1,6]，即所谓的 α 模型。在这个模型中，每次基元分割的几率 α_i 取区间 $[0.5(1-\alpha),\ 0.5(1+\alpha)]$ 中均匀分布的随机数，在此 α 是一个参数，代表涨落的程度。在图 1 a 中画出 3 个不同 α 值时的基元熵 $S(\alpha)$ 的分布。当 $\alpha = 0$ 时，这个分布是在 $S(\alpha) = 0.693$ 处的一个 δ 函数。当 α 增加时，峰的位置偏离 0.693，它的高度也越来越低。注意，对每一个 α 小于 1 的值，基元分割熵 S_α 有一个极小值：

$$S_{\alpha\min} = -[(1-\alpha)\ln(1-\alpha) + (1+\alpha)\ln(1+\alpha) - 2\ln 2]. \tag{13}$$

这是所用的 α 模型的特殊性质，并不是自相似级联机制所必然要求的。

图 2 中的实线代表平均基元熵 $\langle S_\alpha\rangle$。当 $\alpha = 0$ 时，$\langle S_\alpha\rangle = 0.693$，当 α 增加时，$\langle S_\alpha\rangle$ 降低，在 $\alpha = 1$ 时，$\langle S_\alpha\rangle$ 降为 0.5。注意，即使 α 取它的最大值 1 时，$\langle S_\alpha\rangle$ 也不等于(7)

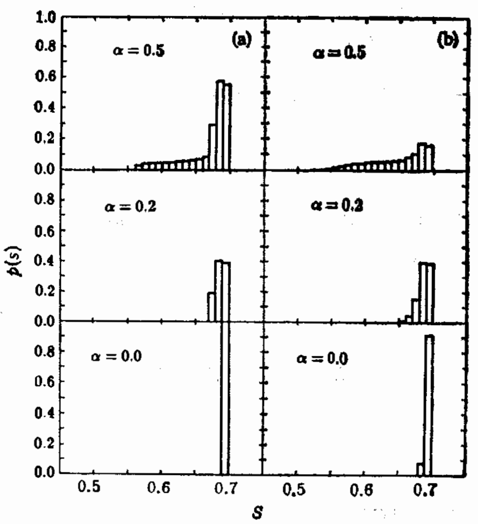

图 1　三种不同 α 值时，基元分割的熵 S_α 的分布。(a) 没有统计涨落。(b) 带有统计涨落，其中 $K = 500$

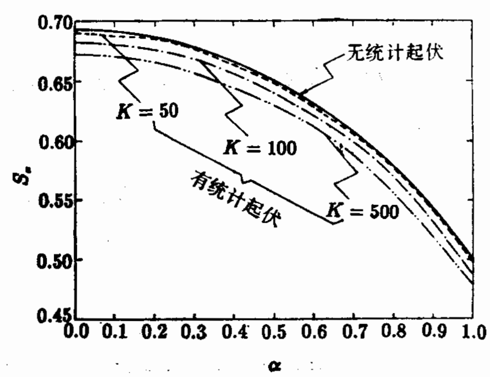

图 2　平均基元熵随 α 的变化。实线是没有统计涨落；双点线，单点线和虚线都是有统计涨落的情况，其中 K 分别为 50, 100 和 500

式所给出的极小值0. 这是由于$\alpha=1$代表α模型中的最大涨落. 在这种情况下，几率a_j并不总是等于1,而是区间[0,1]中的一个随机数.

最后,让我们来讨论如何在分析实验数据中应用上述结果. 在实际的实验中,存在着动力学几率分布基础上的统计涨落. 它使得熵的分布变宽,而平均熵值变小,这一效应在多重数高时变弱. 在多重数足够高时，统计涨落的影响小到可以忽略. 上面给出的讨论近似适用为了说明这点，我们在所给的α模型的基元几率上用二项式分布模拟了粒子的统计涨落. 对多重数 $K=50,100,500$ 的情况模拟得到的结果 $\langle S_a\rangle$ 画在图2中. 为了检验(12)式，计算中用两代的平均熵差作为平均基元分割的熵. 从图可以看出，当$K=500$时，具有统计涨落的平均熵差与纯动力学的情况非常接近. 这意味着对于 $K\geqslant 500$的事件,统计涨落可以忽略,上面的讨论近似适用. 在现有能量下，足够重的核-核碰撞[7]以及在未来的超级加速器,如 SSC 和 Eloisatron[8,9] 上的强子-强子碰撞都可以满足这一条件.

在图1b中,画出了具有统计涨落的样本的基元子分割的几率熵,可以看到分布的形式除了由于归一化引起的一个长而低的尾部外,非常类似于对应的纯α模型的分布(参见图1a) 这进一步肯定了 $K\simeq 500$ 是可以忽略统计涨落并近似应用所讨论结果的下界.

参　考　文　献

[1] A. Bialas and R. Peschanski, *Nucl. Phys.*, **B273** (1986), **703**; *ibid.*, **B308**(1988), 857.
[2] See for example Festschift Léon Van Hove, eds. A. Giovannini and W. Kittel (World Scientific 1990) and the papers cited therein.
[3] See for example R. J. McEliece, The theory of information and coding, Encyclopedia of mathematics and its applications Vol. 3 (Addison-Wesley 1977).
[4] V. S. Pugachev, Probability theory and mathematical statistics for engineers (Pergamon Press 1984).
[5] A. Papoulis, Probability, random variables and stochastic processes (McGraw-Hill 1984).
[6] 吴元芳,刘连寿,科学通报,36卷(1990),21.
[7] G. A. Baym, P. Braun-Munzinger and S. Nagamiya (eds.), Proc. of the Seventh International Conference on Ultra-Relativistic Nucleus-Nucleus Collisions, *Nucl. Phys.*, **A498**(1990).
[8] 吴元芳,刘连寿,高能物理与核物理,**13**(1989),808.
[9] T. D. Lee, Talk given at the Ettore Majorana Center for Scientific Culture, Erice: 28th Course, Physics Up to 200 TeV.

Conditional Entropy and Mutual Information in Random Cascading Processes

WU YUANFANG　　LIU LIANSHOU

(*Institute of Particle Physics, Hua-Zhong Normal University, Wuhan* 430070)

ABSTRACT

The conditional entropy and mutual information of two neighboring generations of division are proposed to use as characteristics in analyzing the random cascading mechanism in multiparticle production. It is shown that these characteristics are effective in examining the exsistence and property of self-similar cascading.