

Use distribution of Hamming distances to determine the phase transitions of the clock model and Generalized xy model.

李宗翰 (S10210028)

指導老師:吳桂光

*Email:s10210028@thu.edu.tw

摘要

clock model 在不同的 q 時會有不同的相變型態，而我嘗試使用 Hamming distances 的分佈機率，經過數據擬和後，來判斷不同 q 以及系統尺寸的 clock model 相變的臨界溫度。

研究計畫內容

(一) 研究動機與研究問題

由於近年來機械學習的技術越來越成熟，我們已經可以透過機械學習來分辨不同相之間的差異，並從中找出臨界溫度。但是這些方法大多都需要先驗知識，因此在面對沒有先驗知識的情況下，計算會變得十分困難。因此我想要了解如果在沒有先驗知識；並且不依賴特定的物理量的條件下，要如何只利用自旋配置的原始數據找出臨界溫度。

而這篇論文[參考文獻1]中的方法，只需要用到資料集中的自旋配置，十分符合上述的條件。因此我在這次研究中，將這個方法套用在了clock model上，在這個過程中學習python的程式語言技巧，以及論文相關的物理知識。

(二) 文獻回顧與探討

在一般的磁性系統中，我們可以利用序參量或是磁化率來判斷它的相變溫度。但是受geometric frustration影响的自旋系統，由於其幾何性質導致不存在唯一的能量最低態，導致科學家在研究相變的過程中出現困難。之後有科學家發現，Hamming distances的分布可以用來檢測受geometric frustration影响的自旋系統的相變，而且不需要先驗知識。[參考文獻1]

這次所使用的clock model被視為xy model的離散版，這個模型的Hamiltonian可以寫成：

$$H_{clock} = - \sum_{\langle i,j \rangle} \cos(\theta_i - \theta_j)$$

而在位置 i 的自旋可以用 $\theta_i = 2\pi n_i/q$ 來表示，其中 $n_i = 0, 1, 2, \dots, q-1$ ，而 $\langle i, j \rangle$ 表示最近鄰的自旋。[參考文獻2]

clock model在 $q < 5$ 的時候為Ising-like相變，只會發生一次相變。而當 $q > 5$ 之後會出現KT相變，會存在兩個不同的相變溫度 T_{c1} 和 T_{c2} ($T_{c2} > T_{c1}$)。當溫度小於 T_{c1} 時為LRO相，而介於 T_{c1} 和 T_{c2} 之間的相為QLRO相，而當溫度大於 T_{c2} 時為無序相。[參考文獻3]

(三) 研究方法及步驟

兩個自旋配置 σ 和 σ' 之間的歸一化Hamming distances定義如下：

$$D(\vec{\sigma}, \vec{\sigma}') = \frac{1}{N} \sum_i \left(\frac{1}{2} - \sigma_i \sigma'_i / 2 \right)$$

由這個公式得出的D，滿足相同的配置D=0，兩個完全相反的配置之間D=1。對於完全不相關的資料集，Hamming distances會以D/N=1/2為中心的高斯分佈。

而以下是這次我所使用的clock model資料集配置：

q=2(L20): 溫度0.5~4，每0.05劃分一個溫度，共70個溫度，每個溫度1000筆資料，共70000筆資料。

q=3(L20): 溫度0~3，每0.05劃分一個溫度，共60個溫度，每個溫度1000筆資料，共60000筆資料。

q=6(L20): 溫度0~1.5，每0.01劃分一個溫度，共150個溫度，每個溫度1000筆資料，共150000筆資料。

q=6(L60): 溫度0.6~2.1，每0.01劃分一個溫度，共150個溫度，每個溫度1000筆資料，共150000筆資料。

接著我畫出了Hamming distances隨著溫度T的分佈圖，在clock model q=2、3的時候，在圖FIG.1中只會看到兩個區域，分別代表LRO相以及無序相，而在低溫的時候Hamming distances可能的值只有2個，這是因為低溫的時候系統為有序相，因此在同一筆資料中其自旋方向幾乎相同，因此兩筆不同資料之間的自旋配置不是完全相同就是完全相反。而根據Hamming distances的定義，相同及相反的自旋配置其D分別等於0和1，因此低溫狀態下，q=2時D的值只會出現0和1。而q=3則因為沒有完全相反的配置， $\sigma_i \sigma'_i$ 所出現的值只能為:1、-0.5，因此D可能出現的值只能為:0、0.75。而隨著溫度逐漸上升，變為無序相之後，自旋的排列開始混亂，因此P(D)=0、1的峰值會逐漸偏向中間(0.5)，並且P(D)會以0.5為中心呈高斯分佈，代表每種自旋配置的機率都相同。

而在clock model q=6的時候可以看出三個不同的區域，分別代表了LRO相、QLRO相、無序相，從FIG.2中可以看出在低溫的時候Hamming distances可能的值有4個。這是因為在q=6的時候，自旋的方向不再是只有完全相同和相反，而 $\sigma_i \sigma'_i$ 所出現的值有可能為:1、0.5、-0.5、-1，因此D可能出現的值為:0、0.25、0.75、1一共四個。而隨著溫度逐漸上升，P(D)=0、1的峰值會逐漸偏向中間(0.5)，此時為QLRO相。當溫度夠高到達無序相時，P(D)會以0.5為中心呈高斯分佈，代表每種自旋配置的機率相同。

FIG.1 跟 FIG.2 中的白色虛線代表了真實的相變溫度，接下來我會利用Hamming distances 分佈的第一個峰值，嘗試去判斷真實的相變點。

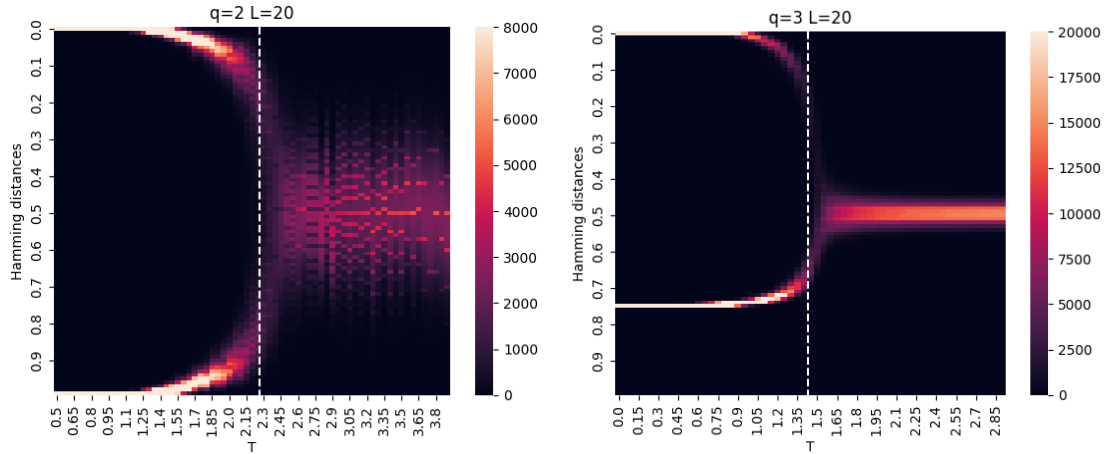


FIG.1. clock model :[q=2、L=20、臨界溫度:2.27]、[q=3、L=20、臨界溫度:1.483]，Hamming distances隨溫度的heatmap

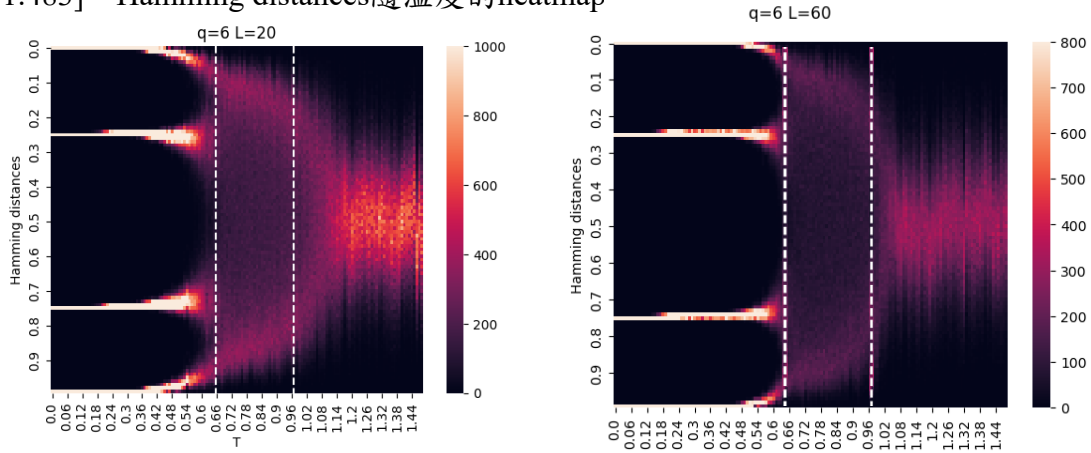


FIG.2. clock model :[q=6、L=20]、[q=6、L=60]、臨界溫度: $T_{c1} = 0.644$ ， $T_{c2} = 0.9605$ ，Hamming distances隨溫度的heatmap。

接著我利用histogram函式將Hamming distances分為100個區域，計算出每個區域內的分佈數量，將其除以總數得出該數值的Hamming distances對應的分布機率 $P(D)$ 。接著利用curve fit函數將不同溫度的 $P(D)$ 分布，對兩個高斯函數的和來進行擬和。而因為 $P(D)$ 分佈對稱於0.5，因此擬和的範圍為 $P(D) \in [0,0.5]$ 即可(為了擬和更加的方便，我選定的範圍為 $[0,0.6]$ ，和論文不同)，而目標函數如下:

$$f(D) = \frac{A}{\Delta\sqrt{2\pi}} \exp\left[-\frac{\left(\frac{D-D_0}{\Delta}\right)^2}{2}\right] + \frac{A'}{\Delta'\sqrt{2\pi}} \exp\left[-\frac{\left(\frac{D-D_0'}{\Delta'}\right)^2}{2}\right]$$

(在q=2、3、範圍 $P(D) \in [0,0.6]$ 時，低溫時只會有一個峰值，因此目標函數為單個高斯函數) 公式為:

$$f(D) = \frac{A}{\Delta\sqrt{2\pi}} \exp\left[-\frac{\left(\frac{D-D_0}{\Delta}\right)^2}{2}\right]$$

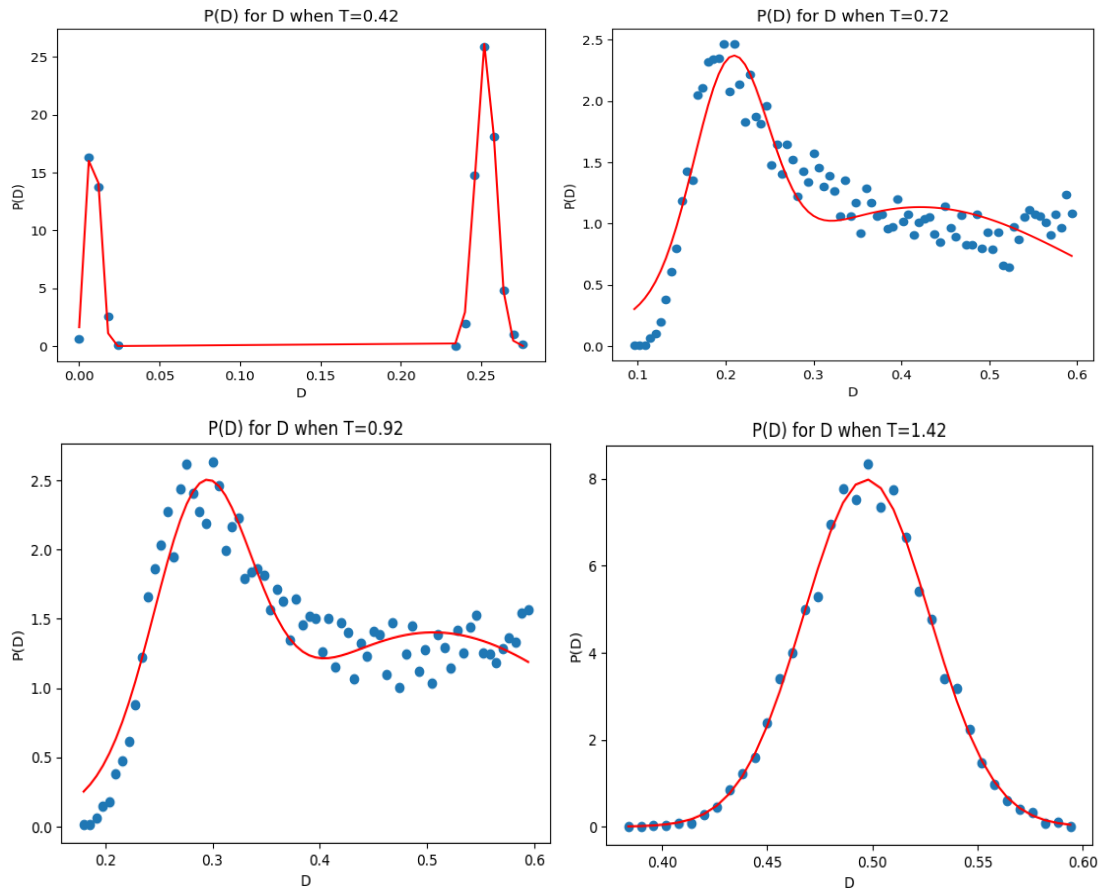


FIG.3.以clock model $q=6$ 、 $L=20$ 為例，在溫度 T 分別為0.42、0.72、1.02、1.42時，curve fit函數對不同溫度的 $P(D)$ 分佈，對兩個高斯函數的和進行擬和的曲線。

FIG.3擬和出來的結果如下：

[$T=0.42$: $D_0 = 0.00868$ 、 $D_0' = 0.25273$]、[$T=0.72$: $D_0 = 0.20598$ 、 $D_0' = 0.42143$]
 [$T=0.92$: $D_0 = 0.29042$ 、 $D_0' = 0.50563$]、[$T=1.42$: $D_0 = 0.4972$]。可以看到隨著溫度升高， $P(D)$ 的峰值會逐漸偏向0.5。要注意的是，在高溫的時候由於Hamming distances的分佈只有一個峰值，此時目標函數需要改為單個高斯函數。因此在溫度 $T > 1.03$ 後我會將目標函數改為單個高斯函數，

接下來將每個不同溫度所對應的第一個峰值(D_0)作圖，從峰值的變化判斷相變點。

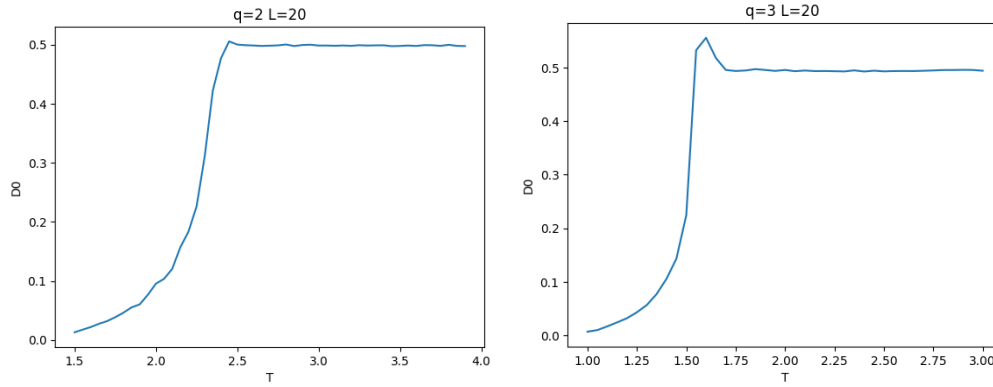


FIG.4. clock model : $[q=2, L=20]$ 、 $[q=3, L=20]$ ， D_0 隨溫度 T 的關係圖。

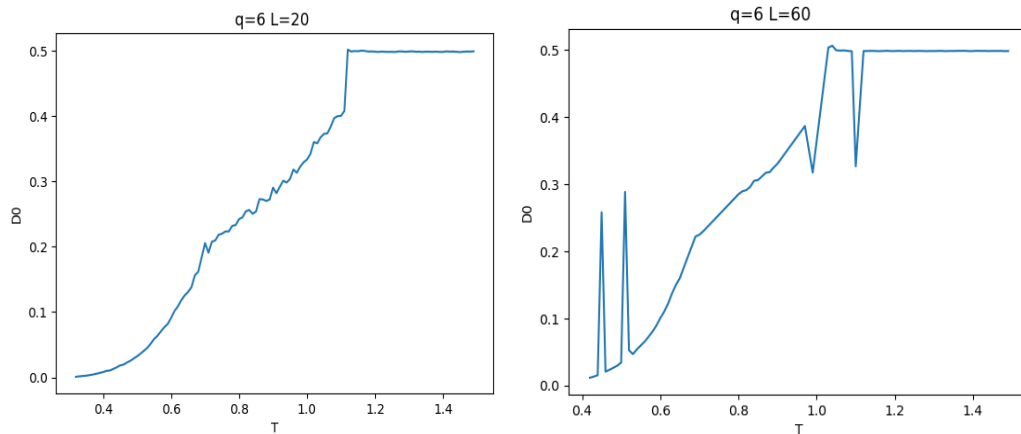


FIG.5. clock model : $[q=6, L=20]$ 、 $[q=6, L=60]$ ， D_0 隨溫度 T 的關係圖。

*其中 $[q=3, L=20]$ 、 $[q=6, L=60]$ 的曲線中間有異常的突起，這是因為在設定擬和參數的時候需要手動設置，而這些異常點是沒有設置到的。而目前發現在擬和時把curve fit函數更改為savgol filter函數可以解決這個問題。

FIG.4和FIG.5中，其斜率變化最大的點所對應的溫度，就代表臨界溫度。可以看到在 $q=2, 3$ 時只會有一個變異點，代表只有一次的相變。而 $q=6$ 的clock model為KT-like相變，有兩個相變點分別位在兩個斜率變化最大時的溫度上。並且兩個不同系統尺寸的 T_{c1} 是相近的，而 T_{c2} 隨著系統尺寸的增加，相變點也會往左移，而更加地接近真實的臨界溫度。

(四) 預期結果

1. 此種方法可以應用在Generalized xy model上，來證明此種方法可以應用在不同的模型上。
2. 利用 D_0 去判斷在 $q=6$ 的clock model的相變點，隨著系統尺寸增加， T_{c1} 不變，而 T_{c2} 會隨著系統尺寸的增加，而往左偏移。
3. 利用有限尺寸縮放來推測相變點。

(五) 參考文獻

- [1] Wen-Yu Su, Feng Hu, Chen Cheng, and Nvsen Ma, Berezinskii-Kosterlitz-Thouless phase transitions in a Kagome spin ice by a quantifying Monte Carlo process: Distribution of Hamming distances, Phys. Rev. B 108, 134422 (2023)

- [2] Kwai-Kong Ng and Min-Fong Yang, Unsupervised learning of phase transitions via modified anomaly detection with autoencoders, *Phys. Rev. B* 108, 214428 (2023)
- [3] S. Chatterjee, S. Puri, and R. Paul, Ordering kinetics in the q-state clock model: Scaling properties and growth laws, *Phys. Rev. E* 98, 032109 (2018)