2 動機實例與分析策略

在本章中我們將以一筆晶圓在薄膜沉積製程中, 化學氣相沉積 (Chemical Vapor Deposition, 簡稱 CVD) 的眞實數據來探討我們的研究動機, 並提出一個創新的資料分析策略, 在本章末亦會說明本研究中的實驗設計課題與傳統設計不同之處。這組眞實數據在晶圓上選取了121個晶粒, 並量測這些晶粒上的薄膜沉積厚度, 以下我們將其視爲反應變數 (response variable)。分析這筆數據的主要目的爲建立一個配適得好的統計模型, 來了解隨著晶粒幾何位置的變化, 薄膜沉積厚度如何改變。

2.1 圓區間與極座標區間

若我們將晶圓以單位圓表示, 並把圓上這121個點的位置描繪出來, 如圖2.1(a) 所示, 我們可以看出在圓上這些點分佈得相當均勻。若觀察這些點上的反應值透過 regres-

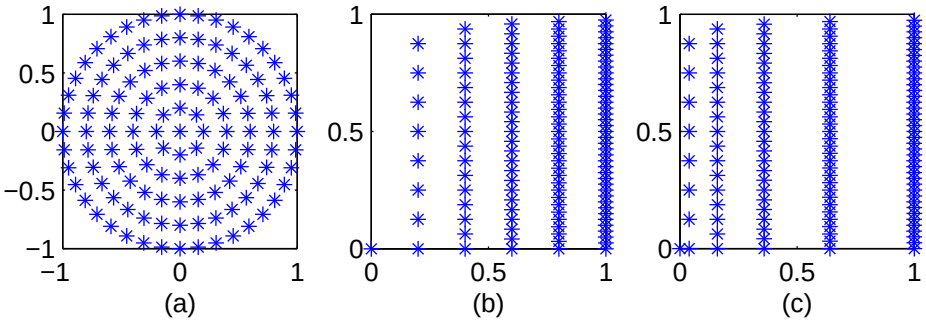

圖 2.1: (a) 實驗區間 (b)(R, θ) 區間 (c)(r, θ) 區間

sion spline 得到的反應曲面圖, 如圖2.2(a) 所示, 我們可發現這組數據有圓效應的存在 , 即這些點的反應曲面主要隨著點到圓心的距離 (R) 改變而變化。由圖2.2(a) 可發現, 這些點的反應曲面一開始隨著 R 增加會遞增, 但是當 R 到某個範圍後會再往外翻。若我們將這些點在圓上的直角座標以 (x, y) 表示, 反應值以 z 表示, 在圓上配適

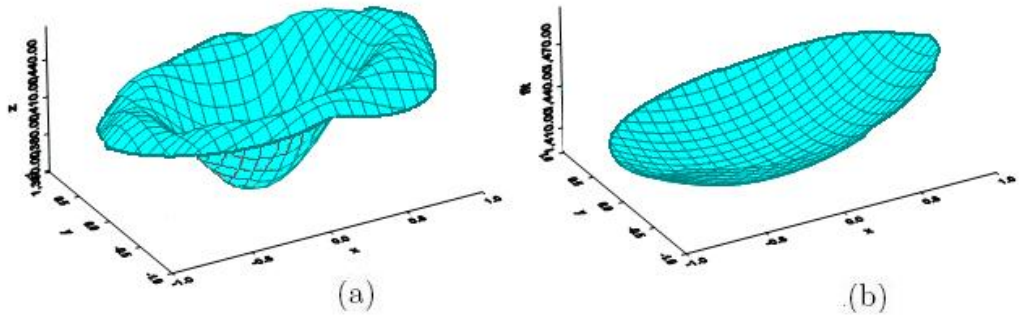

圖 2.2: 在圓上之反應曲面圖

一個二階模型, 則配適模型如下

$$E(z) = 1394.37 + 19.77x - 8.59y + 42.60x^2 + 38.14y^2 - 2.86xy, \tag{2.1}$$

其判定係數 (R-square) 爲0.4481, 故此模型對這組數據的解釋能力並不高 , 圖2.2(b) 爲此二階配適模型的反應曲面圖, 藉由比較圖2.2(a) 與2.2(b), 我們可看出此二階配適模型並未將眞實數據在半徑大到某程度後反應曲面向外翻的特性表現出來。若配適一個三階模型其判定係數雖高一些, 但也只有0.4691。因此對這組數據, 若要在圓上建立模型並且希望其配適得好, 需使用更高階更複雜的模型, 其將需要更多的自由度。

對於具有圓效應的數據, 在極座標區間配適模型, 常會比在圓區間更恰當。在圓上以直角座標表示的點 (x, y), 其極座標值可以透過下列的轉換得到:

$$R = \sqrt{x^2 + y^2},$$

$$\theta = \begin{cases} \cos^{-1}(\frac{x}{\sqrt{x^2+y^2}}), \text{ if } y > 0, \\ 2\pi - \cos^{-1}(\frac{x}{\sqrt{x^2+y^2}}), \text{ if } y \leq 0。 \end{cases}$$

其中 R 爲該點到圓心的距離, θ 爲由 x 軸沿逆時鐘方向算起至該點的角度。若我們將圖 2.1(a) 中的實驗點轉換到極座標區間上, 則這些點在極座標區間的分布如圖 2.1(b) 所示。若我們在圖 2.1(b) 的極座標區間上對 R 與 θ 配適一個二階模型, 則配適模型如下:

$$E(z) = 1321.13 + 355.55R - 170.98\theta - 237.98R^2 + 151.75\theta^2 + 41.02R\theta \quad (2.2)$$

其判定係數爲 0.8924, 故此模型配適得遠優於 (2.1) 式中的模型。若要了解其原因, 我們可將圖 2.2(a) 以極座標區間表示則如圖 2.3(a) 所示。由圖 2.3(a) 我們看出反應值隨著 R 的變動變化較明顯, 而 θ 變化對反應值來說影響沒 R 那麼明顯, 所以 R 效應的影響會比 θ 來得大。圖 2.3(b) 爲 (2.2) 式中的二階配適反應曲面圖。比較圖 2.3(a) 和 (b), 我們可看出此 (R, θ) 區間上的二階配適模型頗能呈現眞實數據的反應曲面之變化。因此我們可得知, 對於具有圓效應的數據, 若要配適一個二階模型, 極座標區間會比圓區間更適合。在極座標區間上只需配適一個簡單模型即可逼近眞實數據的反應曲面, 而在圓區間上則需要一個複雜的高階模型才能做得到。

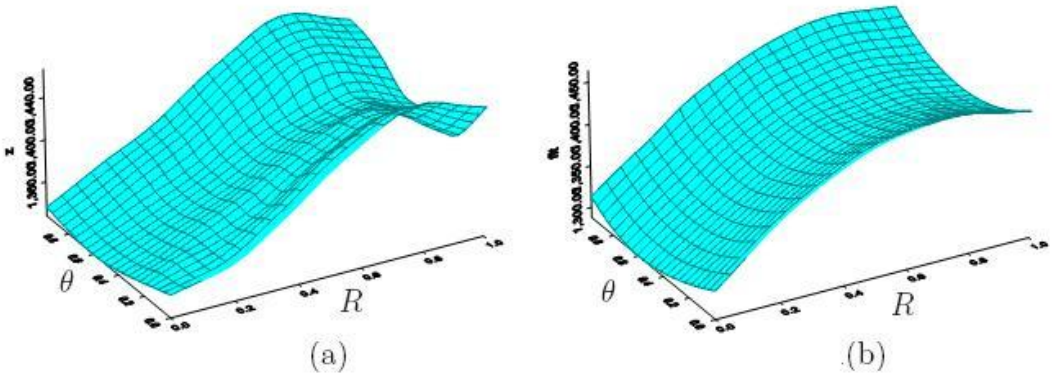

圖 2.3: 在 (R, θ) 區間上之反應曲面圖

由圖 2.1(a) 我們可以發現這 121 個實驗點在圓上的分布很均勻, 但是把這些點轉換到極座標區間後, 由圖 2.1(b) 可看出在 R 大時明顯實驗點較多, 故在 (R, θ) 區間

實驗點的分佈是不均勻的。若我們再將 (R,θ)區間上的點轉換至 (r,θ) 區間上, 其中 $r=R^2$, 如圖2.1(c) 所示, 可以看出實驗點在 (r,θ) 區間上的分布比其在 (R,θ) 區間上的分布更均勻一些。關於圓區間、(R,θ) 區間及 (r,θ) 區間上的實驗點分布的均勻性問題, 我們將會在下一章中詳細討論。

若我們對圖2.1(c) 中的實驗點亦配適一個二階模型, 其結果如下

$$E(z)=1368.98+283.76r-140.37\theta-218.00r^2+142.84\theta^2+17.17r\theta。$$

此模型之判定係數高達 0.9511, 並且所配適出來的反應曲面圖, 如圖2.4(b) 所示, 與圖2.4(a) 中眞實數據的反應曲面圖相比之下, 我們可看出模型配適的相當好。

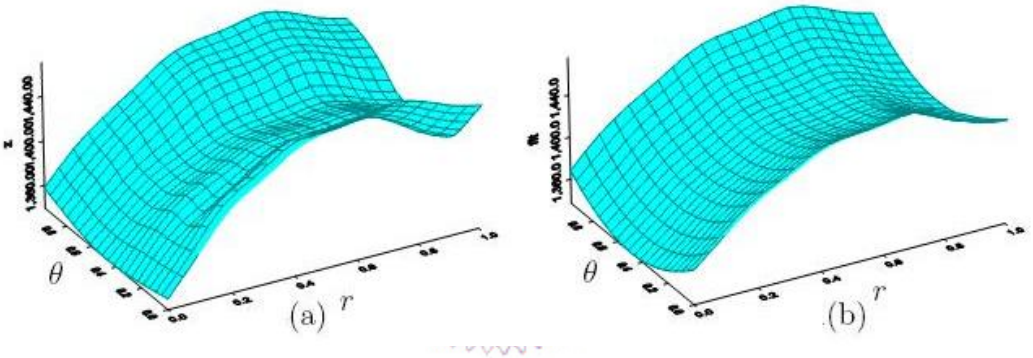

圖 2.4: (r,θ) 區間上之反應曲面圖

2.2 極軸的選取

接下來我們先定義「極軸」。在計算圓上的點的極座標時, 因爲一個圓沒有上下左右的方向性, 因此可以任意決定 $\theta=0$ 所在的位置 (未必需要以 $y=0$ 當成 $\theta=0$), 此時我們定義 $\theta=0$ 所在的位置即爲極軸。假設我們在圓上有 11 個實驗點如圖2.5(a) 所示, 其中粗線表示極軸所在位置。若以 $y=0$ 爲極軸則這些實驗點在 (r,θ) 區間上的分布, 如圖2.5(b) 所示。若改變圓上的極軸位置爲圖2.6(a) 中的粗線, 則以新的極

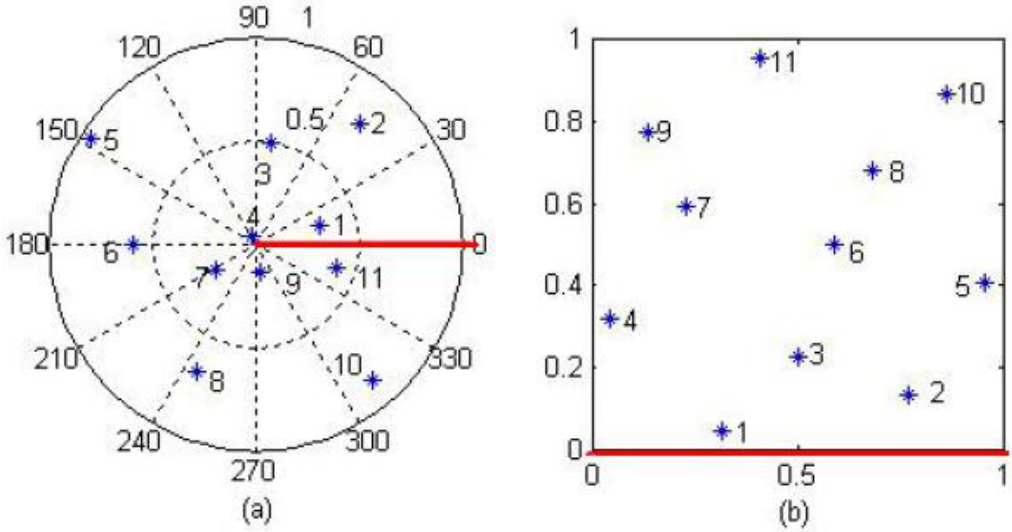

圖 2.5: 未改變極軸之圓區間設計與 (r, θ) 設計

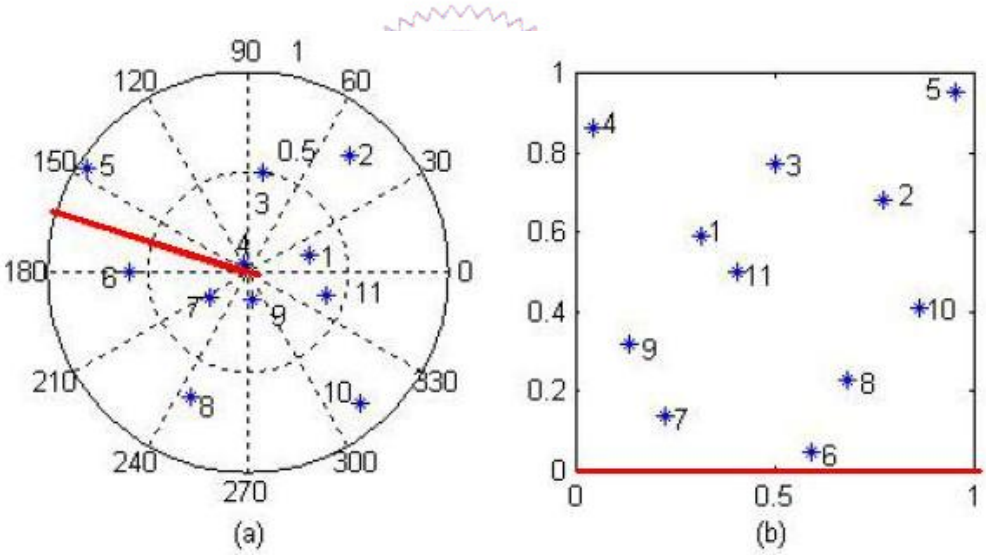

圖 2.6: 極軸改變後之圓區間設計與 (r, θ) 設計

軸展開所對應到 (r, θ) 區間上的實驗點分布如圖2.6(b) 所示。以下我們將說明隨著極軸選取的不同, 圓區間上的實驗點在 (r, θ) 區間上的分布會有什麼變化, 爲簡化陳述, 我們稱 (r, θ) 區間上的實驗點分布爲一個設計。現在我們比較極軸改變前後的 (r, θ)

設計 , 圖2.7(a) 爲圓區間上以 $y = 0$ 爲極軸時產生的 (r, θ) 設計, 當極軸改變爲圖2.6(a) 中的粗線時, 圖2.7(a) 可被此粗線分割爲兩個區塊, 若將下方區塊平移至上方區塊上, 即可得到極軸改變後所對應新的 (r, θ) 設計, 如圖2.7(b) 所示。因此針對同一個圓區間上的設計, 隨著不同的極軸選取, 所得到的 (r, θ) 設計自然不同。改變一次極軸即會對應一個新的 (r, θ) 設計, 但是極軸改變後新的 (r, θ) 設計, 可透過互換原始 (r, θ) 設計的上下區塊之方式來得到。所以在圓上不同極軸的選取, 對應到 (r, θ) 區間上的性質, 即是互換上下區塊。

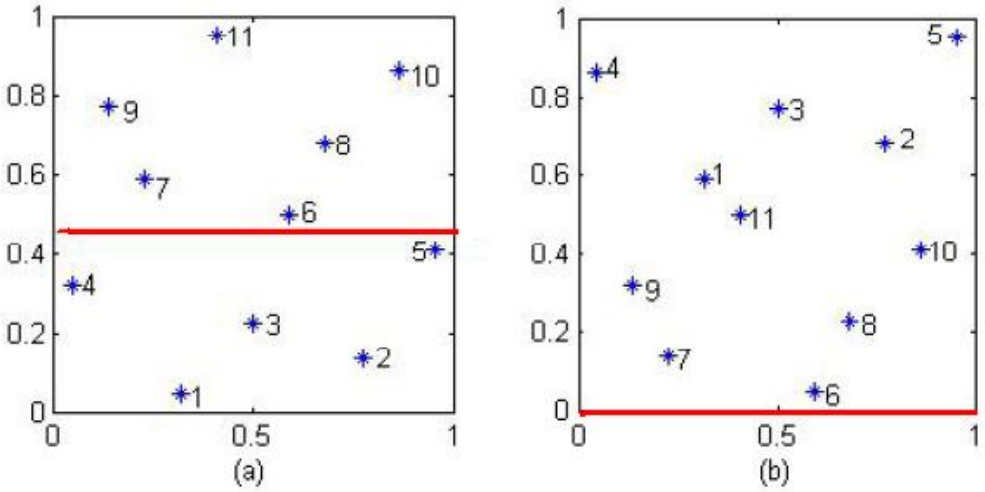

圖 2.7: 不同極軸選取下的 (r, θ) 設計

由於同一個圓區間上的設計, 隨著不同極軸的選取, 所得到的 (r, θ) 設計會不同, 因此我們可知極軸選取對配適模型會有影響。回到 CVD 數據, 以下我們將考慮不同的極軸選取, 對模型配適所產生的影響。在圖2.8中, 我們針對不同的極軸, 各別對其所對應的 (r, θ) 設計, 配適一個二階模型, 並用這些二階配適模型的判定係數對其所相對應的極軸做圖。由此圖我們可以發現, 當極軸通過某個實驗點前後時, 判定係數容易呈現跳動的現象, 但對落在相鄰兩點之間的極軸, 其判定係數不會有太大的差別, 由此圖我們可以清楚看出極軸的選取對配適模型的好壞有影響, 並且隨著極軸選取的不

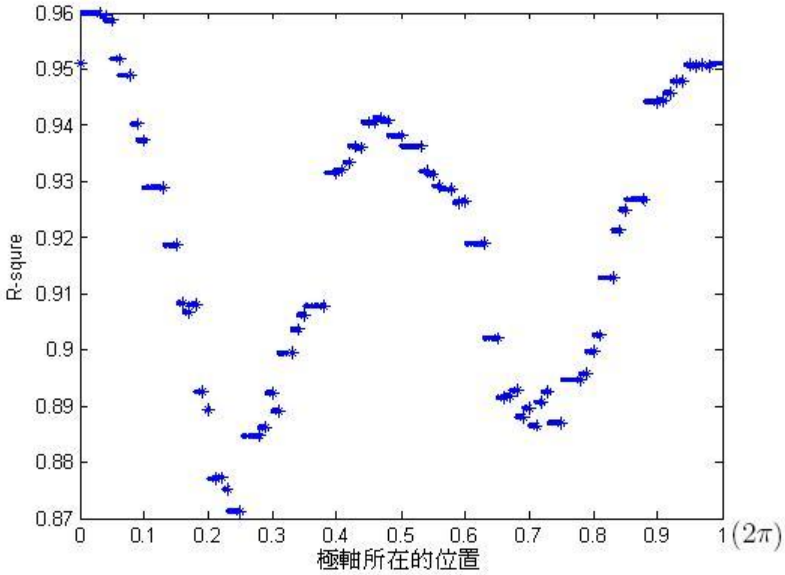

圖 2.8: 不同極軸所在位置與其對應之二階配適模型的判定係數

同, 判定係數可高至 0.9599 或低至 0.8711 。若我們比較最高與最低判定係數所對應 (r, θ) 區間上透過 regression spline 所得到眞實數據的反應曲面圖, 分別如圖2.9(a)、圖2.10(a) 所示, 則可知在圓區間上一組相同的實驗點及其反應值, 隨著極軸選取的不同, 其所導致的 (r, θ) 區間上的反應曲面圖是明顯不同的, 故其二階配適模型的反應

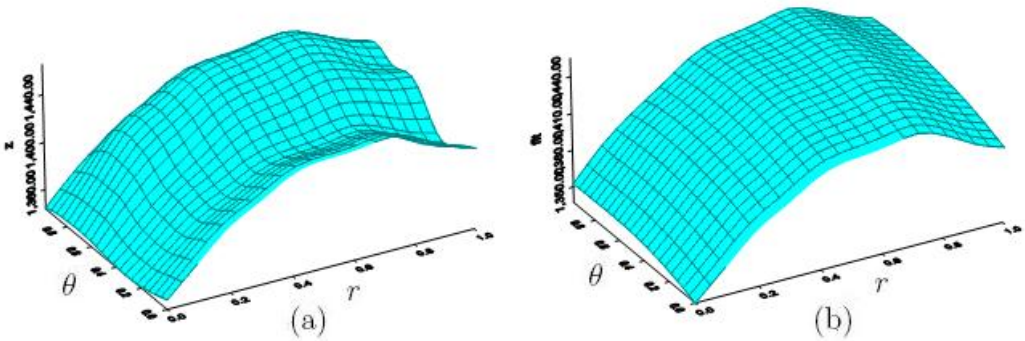

圖 2.9: 極軸選取後配適較差的 (r, θ) 區間上之反應曲面圖

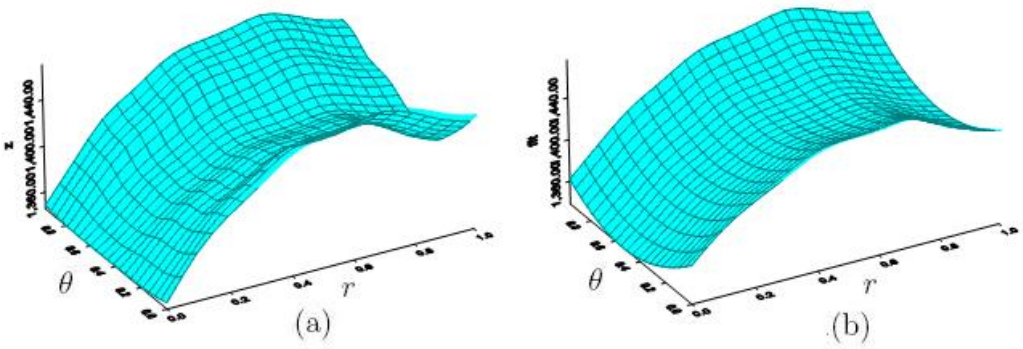

圖 2.10: 極軸選取後配適較好的 (r, θ) 區間上之反應曲面圖

曲面圖 (如圖2.9(b)、圖2.10(b) 所示) 亦會有配適得較好或較差的差別。總而言之, 我們可以挑選極軸來得到一個配適得最好的模型, 因此極軸的選取是分析過程中很重要的一步。

2.3 分析策略

綜合2.1節與2.2節所觀察到的現象, 我們可發現當圓區間上的數據具有圓效應時, 在圓區間上可能需要使用複雜的高階模型才會配適的好, 但在 (r, θ) 區間卻能以簡單的低階模型就將數據的變化描述得好。另外, 由於極軸的選取對模型配適會有影響, 爲了使得模型配適的最好 (判定係數較高), 在分析時我們應改變極軸位置, 來得到最好的配適模型。針對這些現象, 我們對於圓區間上具有圓效應的數據, 提出以下的分析策略:

1 以 (x_i, y_i) 表示第 i 個實驗點在圓區間上的直角座標, 以 z_i 表示第 i 個實驗點上反應變數之觀測值, $i = 1, 2, \cdots, n$。

2 在圓上選取一極軸, 令此極軸與 x 軸正半軸之夾角爲 δ。

3 對此極軸, 計算圓上 n 個實驗點在 (r, θ) 區間的座標值。計算公式如下:

$$r_{\delta,i} = x_i^2 + y_i^2,$$

$$\theta_{\delta,i} = \begin{cases} \cos^{-1}(\frac{x_i}{\sqrt{x_i^2+y_i^2}}) - \delta \quad (mod\ 2\pi), \text{ if } y_i > 0, \\ 2\pi - \cos^{-1}(\frac{x_i}{\sqrt{x_i^2+y_i^2}}) - \delta \quad (mod\ 2\pi), \text{ if } y_i \leq 0。 \end{cases}$$

4 以 $(r_{\delta,i}, \theta_{\delta,i})$ 當解釋變數, 對 z_i 配適一個反應曲面模型(比如二階模型)。令此配適模型爲 $\mathcal{M}_\delta$, 其判定係數爲 $\mathcal{R}_\delta^2$。

5 對 $\delta \in [0, 2\pi]$, 重覆步驟 $2 \sim 4$, 令 δ^* 爲極大化(maximize) $\mathcal{R}_\delta^2$ 的 δ 值, 則 $\mathcal{M}_{\delta^*}$ 爲最終配適模型。

在步驟5中, 雖然應該對 $[0, 2\pi]$ 中所有的 δ 值重覆步驟 $2 \sim 4$ 的計算, 但因對於落在相鄰兩點間的極軸, $\mathcal{R}_\delta^2$ 的值並不會有太大的變化 (見圖2.8), 故在實務上, 我們可以僅在任意相鄰兩點間選取一個 δ 值即可。故步驟 $2 \sim 4$ 僅需重覆 n 次。

2.4 與傳統設計不同之處

當因現實考量無法量測晶圓上所有的晶粒, 而僅能挑選某些晶粒來量測時, 該挑選哪些位置的晶粒, 便是晶圓品管上一個重要的實驗設計問題。以下我們將圓區間上挑選出來的一組實驗點集合稱爲一個設計, 在2.3節的分析策略下, 本研究中的設計與傳統的設計有以下幾個較重要的不同之處。

在傳統的實驗設計中, 實驗區間通常都是長方形區間, 且模型就建立在實驗區間上。但在此我們所要面對的眞實數據的實驗區間是圓形 (以單位圓表示), 而建模時所使用的 (r, θ) 區間則爲正方形 (以單位正方形表示)。故在本研究中, 我們將明確區分實驗區間與建模區間的不同, 實驗區間爲做實驗時所使用的區間, 而建模區間則爲分析時所使用的區間, 在傳統實驗設計中, 通常這兩者是相同的。

此外對於同一個圓區間上的設計, 隨著不同的極軸選取, 將導致 (r, θ) 區間上會出現許多不同的設計。因爲在實驗設計的階段, 反應值尚未被觀測到, 故在未來做分析

時, 該選取哪一個極軸才能得到最好的配適模型, 在設計的階段並無法預測。所以我們必須在設計時要求這些 (r, θ) 區間上的不同設計都必須要有穩定 (robust) 的表現。故我們目前探討的問題必須要求很多組 (r, θ) 區間上的設計都有好的性質, 這與傳統設計僅考慮如何最優化一組實驗點也是截然不同的。

此外, 在傳統因子設計中有許多公認的準則, 如均勻性 (uniformity)、正交性 (orthogonality)、有效性 (efficiency) 等。傳統的實驗設計並沒有區分這些準則是要加於實驗區間上或是建模區間上, 但因在我們考慮的問題中, 實驗與建模區間並不相同, 故我們在這兩個區間上要求該滿足的準則, 亦不相同。比如在此我們希望在實驗區間 (圓) 上的實驗點應均勻散佈, 並且在建模區間 (r, θ) 上配適的模型應有正交的性質。前者可使得圓上不同區域內晶粒的資訊皆能被呈現出來, 而後者則可降低分析時效應之間彼此的影響。

TSING HUA UNIVERSITY