

國立中央大學

數學研究所

碩士論文

大腦視丘神經元之連發波訊號
的自動偵測與分類

Automatic Detection and Classification
of Bursts in Brain Thalamus Neurons

研究生：歐靜文
指導教授：單維彰博士

中華民國九十九年六月



國立中央大學圖書館 碩博士論文電子檔授權書

(98 年 4 月最新修正版)

本授權書所授權之論文全文電子檔(不包含紙本、詳備註 1 說明)，為本人於國立中央大學，撰寫之碩/博士學位論文。(以下請擇一勾選)

- 同意** (立即開放)
- 同意** (一年後開放)，原因是：_____
- 同意** (二年後開放)，原因是：_____
- 同意** (三年後開放)，原因是：_____
- 不同意**，原因是：_____

以非專屬、無償授權國立中央大學圖書館與國家圖書館，基於推動「資源共享、互惠合作」之理念，於回饋社會與學術研究之目的，得不限地域、時間與次數，以紙本、微縮、光碟及其它各種方法將上列論文收錄、重製、公開陳列、與發行，或再授權他人以各種方法重製與利用，並得將數位化之上列論文與論文電子檔以上載網路方式，提供讀者基於個人非營利性質之線上檢索、閱覽、下載或列印。

研究生簽名： 歐靜文 學號： 972201028

論文名稱：大腦視丘神經元之連發波訊號的自動偵測與分類

指導教授姓名：單維彰

系所：數學 所 博士班 碩士班

日期：民國 99 年 6 月 24 日

備註：

1. 本授權書之授權範圍僅限電子檔，紙本論文部分依著作權法第 15 條第 3 款之規定，採推定原則即預設同意圖書館得公開上架閱覽，如您有申請專利或投稿等考量，不同意紙本上架陳列，須另行加填聲明書，詳細說明與紙本聲明書請至 <http://thesis.lib.ncu.edu.tw/> 下載。
2. 本授權書請填寫並親筆簽名後，裝訂於各紙本論文封面後之次頁（全文電子檔內之授權書簽名，可用電腦打字代替）。
3. 請加印一份單張之授權書，填寫並親筆簽名後，於辦理離校時交圖書館（以統一代轉寄給國家圖書館）。
4. 讀者基於個人非營利性質之線上檢索、閱覽、下載或列印上列論文，應依著作權法相關規定辦理。

國立中央大學碩士班研究生

論文指導教授推薦書

數學 學系/研究所 歐靜文 研究生

所提之論文

大腦視丘神經元之連發波訊號的自動偵測與分類

係由本人指導撰述，同意提付審查。

指導教授 單維新 (簽章)

99年5月31日

國立中央大學碩士班研究生
論文口試委員審定書

數學 學系/研究所 歐靜文 研究

生所提之論文

大腦視丘神經元之連發波訊號的自動偵測與分類

經本委員會審議，認定符合碩士資格標準。

學位考試委員會召集人
委員

嚴健彰

蔡亞利

吳維彰

中華民國 99 年 6 月 24 日

大腦視丘神經元之連發波訊號 的自動偵測與分類

學生：歐靜文

指導教授：單維彰博士

國立中央大學數學研究所

摘 要

在神經訊號中，最常見到的是單一棘波訊號 (spikes)，而生物學家在大腦視丘裡紀錄到另一種與 spikes 不同的訊號，稱之為連發波訊號 (bursts)。本文主要在探討 bursts 訊號的特徵，並將特徵具體量化，以利於自動偵測以及分類 bursts 訊號。

首先我們先介紹神經細胞如何產生動作電位，以及神經訊號的相關背景知識。然後，根據生物學家憑經驗挑選出來的 bursts 訊號，我們歸納其特徵並建立了三個篩選條件。這三個篩選條件分別是：與時間相關的「間隔條件」、與振幅相關的「遞減條件」、以及與波形相關的「形似條件」。有了篩選條件之後，接著將這些條件實際運用到一筆原始訊號中去偵測 bursts。在探測訊號之前，會先對原始訊號做一些基本的前置處理，包含降低取樣頻率以及濾波。前置處理完後，利用上述三個篩選條件來篩選 bursts 訊號，並利用主成份分析 (簡稱 PCA) 為 bursts 訊號做分類。

用生物學家憑經驗挑選出來的 bursts 訊號做測試，經由篩選條件過濾後的結果是被生物學家所接受的。而篩選條件自原始訊號中所偵測到的 bursts 訊號，也通過了生物學家的檢驗。由此可以看到篩選條件是能夠反應 bursts 訊號的特徵，並有效的偵測出 bursts 訊號。

Automatic Detection and Classification of Bursts in Brain Thalamus Neurons

Student: Jing-Wen Ou

Advisor: Wei-Chang Shann

Submitted to Department of Mathematics
National Central University

ABSTRACT

The most common signals found in nerve signals are isolated spikes. However, biologists have detected signals other than spikes in the thalamus. These are known as bursts. This paper will mainly be examining the characteristics of bursts. Moreover, the data found will be quantified in order to automatically detect and categorize bursts.

First, we will introduce how nerve cells generate action potentials and the background information of nerve signals. Next, according to the bursts selected by a biologist through experience, we will generalize the characteristics and establish three screening conditions. These three screening conditions are as follows: associated with time, *the gap condition*, associated with amplitude, *the decay condition*, and associated with waveforms, *the shape condition*. With these conditions, we will then apply them on to a set of raw data to detect bursts. Before detecting the signals, we will first process the raw data. This includes down sampling and filtering. After processing the raw data, we will automatically detect the bursts using the three screening conditions mentioned above. In addition, using the Principal Component Analysis (PCA), we will then classify the bursts.

Testing with the bursts that the biologist has selected based on his experience, the results collected through the filtering is confirmed. Moreover, the bursts detected from the raw data using the filtering criteria also pass the test. Therefore, we can see that the filtering criteria can reflect the characteristics of bursts and effectively detect them.

致 謝

兩年的研究所生涯到了這裡，暫時先劃個逗點，

首先，感謝指導老師單維彰教授在這段時間的耐心教導與聆聽，沒有老師的幫忙就不會有這一篇論文。同時更感謝單老師在這兩年於各方面的建議，無論是課業上、生活上或是未來規劃，都給予許多中肯和實用的意見。在每一次的 meeting 中，除了跟老師學習到實用的數學以及計算機知識外，更難得的是從 meeting 裡穿插的小故事中，都獲得了一些科學人文的洗禮，我想，這些小故事可能會成爲我最難忘的事情了。另外還要感謝蔡孟利教授以及嚴健彰教授對於這篇論文的建議和指導。

其次，感謝吳維漢老師，謝謝老師給了我一個訓練口才的機會，讓一向害怕說話且不善言辭的我，總算能夠在大庭廣眾下說出一點點人話了。雖然老師不是我的指導教授，但是我很感謝您！（不過還是很畏懼您）

再次，感謝我的家人，沒有家人的支持，我大概不能毫無顧慮的完成學業。謝謝爸爸讓我了解什麼叫「現實」，謝謝媽媽讓我看到什麼叫「理想」，謝謝我的妹妹們容忍一位經常嘻皮笑臉又不正經的姐姐，不過這兩年在這方面上我沒有什麼進步，未來也不打算改進它。

最後，感謝我的朋友們，如果沒有你們的陪伴，這兩年一定會少了很多歡樂。也許在這裡我沒有辦法一一的點出人名，不過相信你們會收到我由衷的感謝。

現在，我可以劃下句點了。

目 錄

表目錄	ix
圖目錄	x
1 研究動機	1
2 背景知識簡介	3
3 實驗訊號處理與分析	8
3.1 訊號的前置處理	8
3.2 分析訊號與建立篩選條件	12
3.2.1 間隔條件	12
3.2.2 遞減條件	14
3.2.3 形似條件	17
3.3 篩選條件的測試	18
3.4 處理原始訊號以及 bursts 的分類	21
4 振幅衰減的統計模型	27
4.1 數值模擬一： $Y = RX$	28
4.2 數值模擬二： $\mu_Y = r\mu_X$	31
4.3 數值模擬三： $Y = rX + N$	33
4.4 數據模擬結果與真實數據之比較	35
5 結論與展望	37
5.1 結論與探討	37
5.2 檢討與建議	38

參考文獻	41
附錄	42
A.1 降維函式	42
A.2 去雜訊函式	43
A.3 找局部極值函式	44
A.4 偵測 spikes 與 bursts 函式	45
A.5 檢查 bursts 的相對誤差函式	50
A.6 切割 bursts 為 spikes 回傳相對誤差以及代表波形函式	52
A.7 PCA 分析以及 K-means 分群	56
A.8 主程式—處理流程	58

表 目 錄

3.1	前三個 units 的振幅比率表	15
3.2	由三個 units 組成 burst 的 units 振幅比率表	16
3.3	由四個 units 組成 burst 的 units 振幅比率表	16
3.4	篩選條件於不同設定之結果	21
4.1	前二個 units 振幅平均值與標準差	36

圖 目 錄

1.1	spikes 訊號圖	1
1.2	bursts 訊號圖	2
2.1	動作電位各階段示意圖	3
2.2	細胞膜離子變化過程	4
2.3	神經訊號	5
2.4	發生於 8.814~8.82 秒之間的棘波 (spike) 訊號	6
2.5	Bursts 訊號	7
3.1	將 $\sin(x)$ 由 100 個取樣點 down sampling 至 20 個取樣點	9
3.2	白雜訊模擬圖	10
3.3	神經訊號雜訊組成示意圖	10
3.4	實驗中紀錄的神經元電位訊號圖	11
3.5	實驗中紀錄的神經元電位訊號使用 shrink 濾波去雜訊的過程圖	12
3.6	bursts 訊號圖	13
3.7	bursts 平均間隔時間直方圖	13
3.8	bursts 振幅遞減示意圖	14
3.9	迴歸直線斜率與第一個 unit 振幅散佈圖	15
3.10	計算組成 bursts 的 units 波形相對誤差過程圖	18
3.11	Bursts 的相對誤差統計直方圖	19
3.12	112 個 bursts 中振幅未遞減的 bursts 訊號	20

3.13	units 振幅比率直方圖	21
3.14	設定 0.7 為探測動作電位的門檻值	22
3.15	通過第一個篩選條件的錯誤訊號	22
3.16	通過間隔條件以及遮減條件的訊號	23
3.17	通過間隔條件以及遮減條件的錯誤訊號	23
3.18	Bursts PCA 分類圖	24
3.19	Spike (動作電位) 利用 PCA 與 k-meas 分類結果圖	24
3.20	四群動作電位的代表波形	25
3.21	Bursts 波形與動作電位波形比對結果	25
3.22	Bursts 波形與動作電位波形比較	26
4.1	散佈圖	28
4.2	XY 分佈圖	29
4.3	各種常態分佈的統計直方圖	31
4.4	模擬方式二的 XY 散佈圖	32
4.5	相同分佈的常態分佈直方圖	33
4.6	模擬方式三的 XY 散佈圖 I	34
4.7	模擬方式三的 XY 散佈圖 II	35
4.8	雜訊選取	35
4.9	真實數據中的統計模型相關數值表	36
5.1	bursts 振幅間比直直方圖	38
5.2	去雜訊比例與振幅比較圖	39

5.3 動作電位波形示意圖	40
---------------------	----

第 1 章 研究動機

大腦的重量與體積雖然只佔了人體的一小部份，卻主控了整個人的運作，其地位十分重要！瞭解大腦的運作便是生物學上的一個重要主題。

腦部是神經系統的中樞，有相當多的神經細胞聚集在此處。為了瞭解大腦的運作機制，其中一種廣為生物學家所使用方式，是利用微電極探針來紀錄神經細胞在傳遞訊息時膜電位的變化。如果一根探針附近有一個以上的神經細胞，那麼該探針就會紀錄到一種以上的波形。

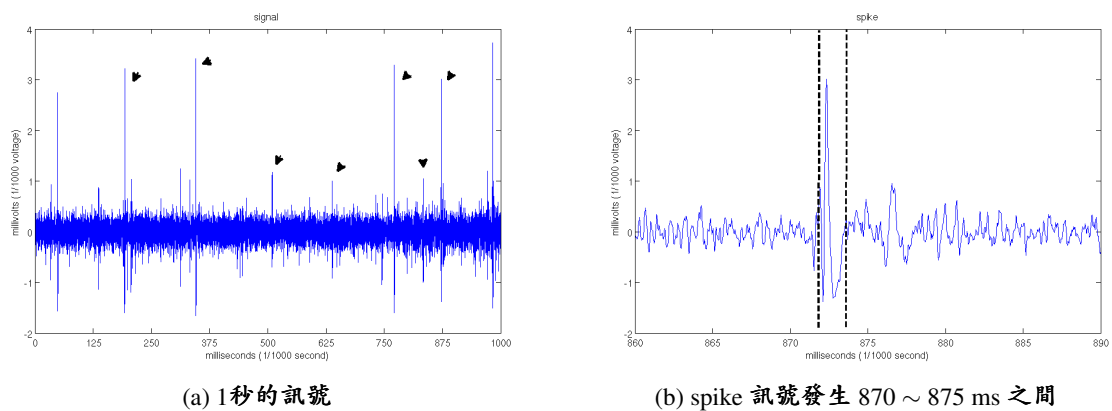


Figure 1.1: spikes 訊號圖

Figure 1.1 是由真實訊號所截取出來的一小部份，從 (a) 圖中可以看到，在某些時間點會有單一明顯的電位變化，只要電位超過雜訊的標準就認為是一次性的膜電位變化，這種一個類似突起般的變化生物學家稱之為棘波 (spikes) 訊號，例如在 (a) 圖中加上箭頭標籤的那些動作電位。我們從 Figure 1.1 (a) 中發生在大約 875 ms (millisecond) 處的 spike 訊號放大至 (b) 圖一般而言，探針在大腦皮質所紀錄到的訊號大多數會如 (b) 圖所示。為了方便起見，在本文中往後將 Figure (b) 此種長像的訊號稱作 spikes。

人受到外界的刺激會有所感覺，例如被針刺到會痛，神經細胞會將這些感覺的訊息先傳送到視丘 (Thalamus)，再送到大腦皮質。可以將視丘想像成一個大型的中繼站，所有感覺訊息 (除了嗅覺外) 在到達大腦前，都會先到視丘。探針在視丘這個訊息轉運站中偵測到兩種不同的訊號模式 [1]，一種是像 Figure 1.1 (b) 圖的 spikes 訊號，另一種則像 Figure 1.2 (b) 。

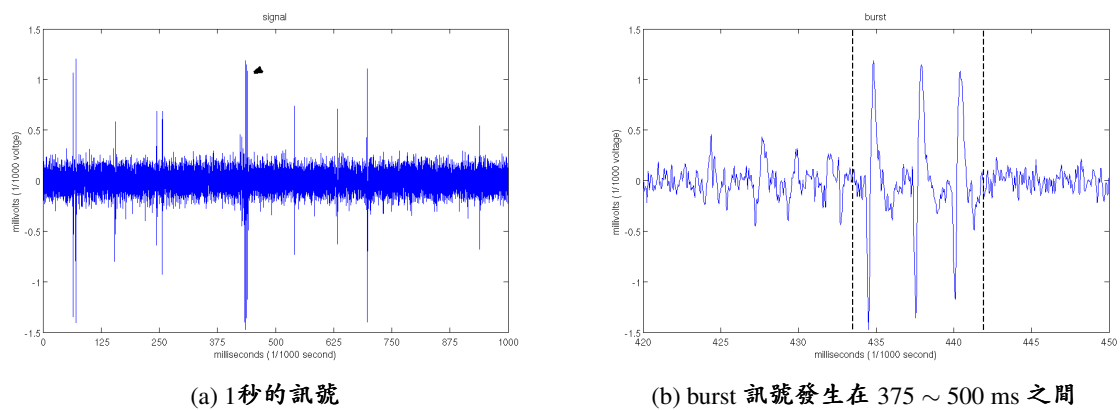


Figure 1.2: bursts 訊號圖

我們看到 Figure 1.2，在 (a) 圖中可以明顯看到大部份的 spikes 都是單一的，但有一些 spikes 卻非常緊密的並在一起，如 (a) 圖箭頭標籤處。這種一次有數個看起來像 spikes 且緊密接連在一起的訊號，生物學家稱之為連發波 (bursts) 訊號。緊密發生的 spikes，有可能單純的恰好是發生時間很接近的 spikes，而也有可能是生物學家所說的 burst 訊號。將 Figure 1.2 (a) 中發生在 375 ~ 500 ms 的 bursts 放大至 (b) 圖，從圖中可以看到，這一個 bursts 是由三個看起來像 spikes 的訊號所組成，而且 spikes 的振幅逐次遞減。

看到了 bursts 具有振幅遞減，且接連發生 spikes 這樣特別的表徵，首先想要討論的事情是，該如何明確地描述 bursts 的這些特性？換句話說，我們想要知道是否可以量化 bursts 的特徵。如果可以量化 bursts 的特徵，那麼便可以透過適當的條件，利用軟體從整筆紀錄中找到發生訊號的時間點 [5] 以及訊號種類。生物學家做實驗時如果可以即時判斷正在活躍的神經細胞個數、細胞產生的訊號種類以及發射訊號的頻率等資訊，就可以對照生物的行為，更瞭解神經系統的變化。

因此，本文主要的目的是分析 bursts 這種特殊形式的訊號，希望能夠根據訊號的特徵設定合理且明確的過濾條件，未來以軟體在探測 (detection) 以及分類 (sorting) 上能夠有所依據。後續幾章內容編排如下：第二章將簡單介紹相關的神經訊號背景知識，若想知道更詳細的內容，請參考 [6]。第三章將描述使用到的相關數學或統計觀念，以及數值模擬。第四章說明訊號處理的過程。第五章為總結與檢討。附錄是本文在實驗中使用到的相關函式，以及函式的使用範例和說明。

第 2 章 背景知識簡介

大腦是人體重要的中樞神經系統，其中包含了上億個神經細胞。眾多的神經細胞組成複雜的神經系統傳遞與分析大量的訊息，使得人類具有認知、學習、記憶...等，這些都是大腦的產物。因此，欲研究大腦的活動，就需要瞭解神經細胞如何傳遞訊息。

細胞膜是一種脂質，因此在正常情況下，多數的脂溶性分子是可以自由地進出細胞，但是帶電離子則不然。帶電離子進出細胞，必須經由細胞膜上的離子通道 (ion channel) 來進出，離子通道上有閘門 (gate)，當閘門開啓時，帶電離子便可透過通道進出細胞。細胞內外的離子濃度並不相同，細胞內的鉀離子濃度較高，鈉離子濃度較低；細胞外的鉀離子濃度較低，鈉離子濃度較高。鈉鉀幫浦 (The Sodium-Potassium Pump) 每作用一次會運送兩個鉀離子進入細胞內，送出三個鈉離子到細胞外。鈉鉀幫浦送出去的正電離子比帶進來的少，有助於維持細胞膜內外的電位差，以及鉀、鈉離子的濃度。

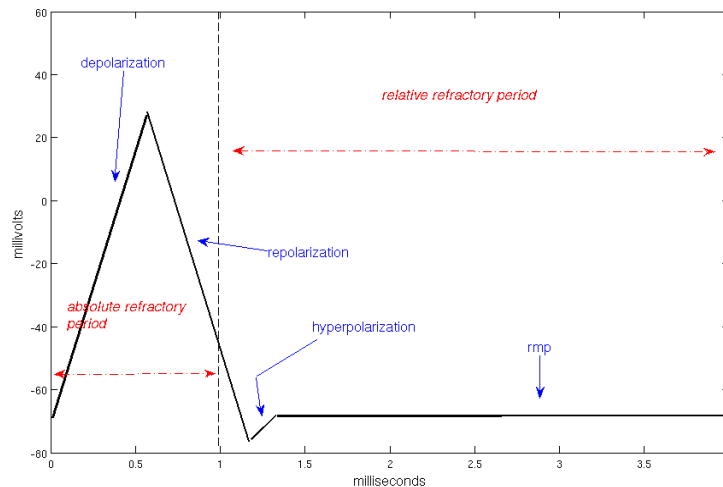


Figure 2.1: 動作電位各階段示意圖

細胞內通常帶有較多的負電，造成細胞膜內外有電位差 (potential difference)。若設定細胞外的電位為 0，一般而言，神經細胞在靜止 (也就是未受到刺激) 時，細胞膜內電位約為 -70 mV (millivolts)，此稱為靜止膜電位 (Resting membrane potential 或簡稱 rmp)。在靜止時，鈉離子通道以及鉀離子通道皆為關閉狀態，當神經細胞受到刺激後，鈉離子通道將打開，使得大量的鈉離子經由鈉離子通道擴散

至細胞內，讓細胞內的電位差快速上升，改變了細胞膜內外的電位差，這個電位改變的過程稱為去極化（Depolarization）。去極化破壞了細胞膜內外的平衡電位，當去極化達到某種程度時，鈉離子通道關閉，而鉀離子通道打開，使得大量的鉀離子經由鉀離子通道擴散至細胞外，讓細胞內的電位差下降，這個過程稱之為再極化（Repolarization）。再極化關閉了鈉離子通道打開鉀離子通道，此時因為細胞內外的鉀離子濃度不均，使得鉀離子不斷地向外擴散，造成細胞內的電位差比靜止時的電位差更負，這個過程稱之為過極化（Hyperpolarization）。神經細胞受到刺激後的去極化、再極化以及過極化這一連串的過程，稱之為動作電位（Action Potentials）或是神經衝動（Nerve impulse）。

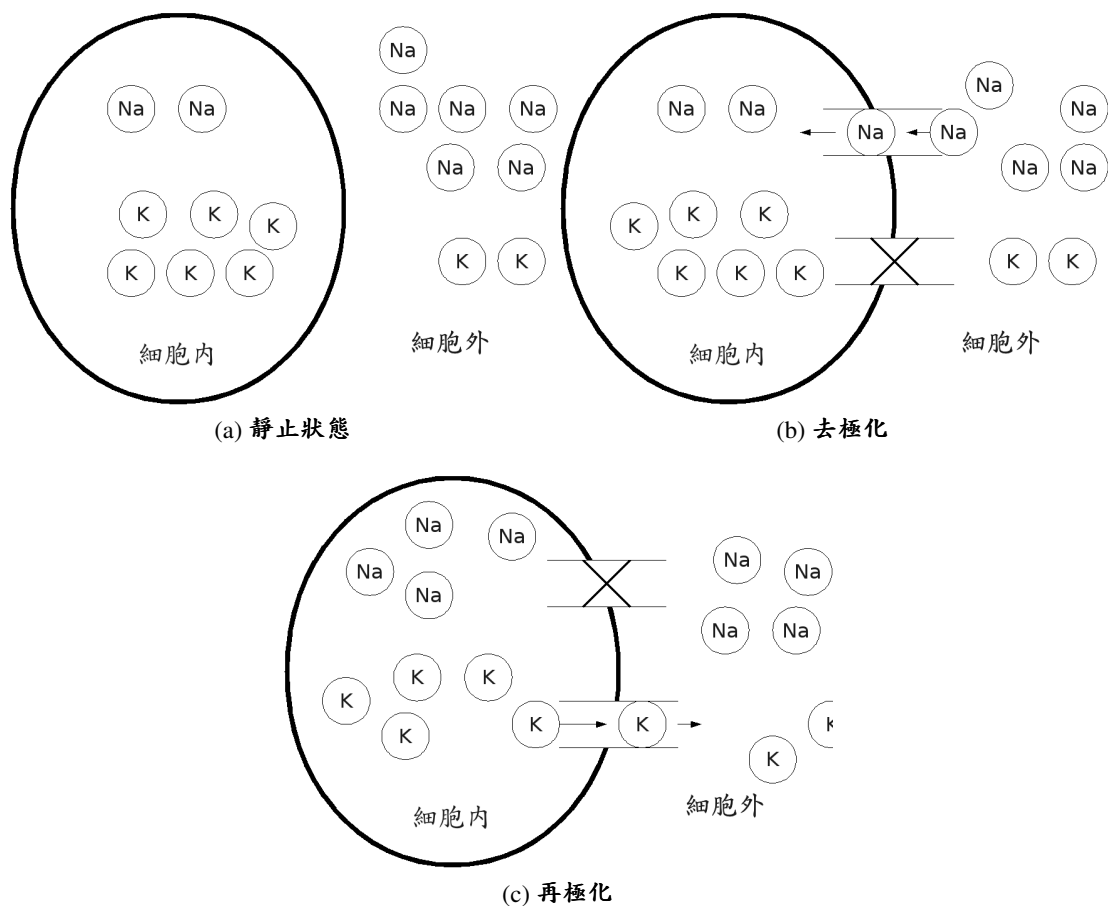


Figure 2.2: 細胞膜離子變化過程

當神經細胞受到刺激而去極化到達某個閾值時，細胞將產生動作電位，不達閾值即沒有反應。只要超過了閾值，不管是大的刺激，或是小的刺激，其產生的動作電位幅度都相等，大的刺激會產生較多的動作電位，小的刺激產生較少的動作電位。也就是說，神經細胞是以動作電位產生的頻率來表示刺激的強弱，不是以大小

來表示。換句話說，同一個神經細胞，無論所受到刺激是強或者弱，其產生的動作電位大小都相等，這種現象稱之為全有全無律（all-or-none law）。

當神經細胞正在發生動作電位的初期，若再給予第二個刺激，細胞並不會對第二個刺激產生反應，這段不反應的時間稱為絕對不反應期（absolute refractory period）。也就是說，在絕對不反應期間，無論給予再大再強的刺激，神經細胞都不會再產生動作電位。絕對不反應期通常為 1 毫秒，也就是 0.001 秒。在動作電位的後期，也就是膜電位回復到接近閾值的狀態時，若此時給予相當強烈的刺激，那麼將有可能再產生動作電位。這段時間稱為相對不反應期（relative refractory period）。相對不反應期通常為 4 毫秒，也就是 0.004 秒。舉個簡單的例子：假設一個神經細胞 A 在正常情況之下，產生動作電位的門檻值是 2mV（millivolts）。A 細胞在第 x 毫秒時受到 2.1mV 的刺激後產生動作電位，在 $x \sim x + 1$ 毫秒之間若同樣給予 2.1mV，甚至 200mV 的刺激，A 細胞都不會再次產生動作電位， $x \sim x + 1$ 毫秒這一段時間為絕對不反應期。過了絕對不反應期後，在 $x + 1 \sim x + 4$ 毫秒間，若給予的刺激是 2.1mV，則 A 細胞仍不會產生動作電位；但若給予的刺激遠大於門檻值，例如 200 mV，則 A 細胞將產生第二個動作電位， $x + 1 \sim x + 4$ 毫秒這一段時間為相對不反應期。過了相對不反應期，也就是在 $x + 4$ 毫秒之後，A 細胞就回復到了正常情況，此時給予 2.1 mV 的刺激即可產生動作電位。

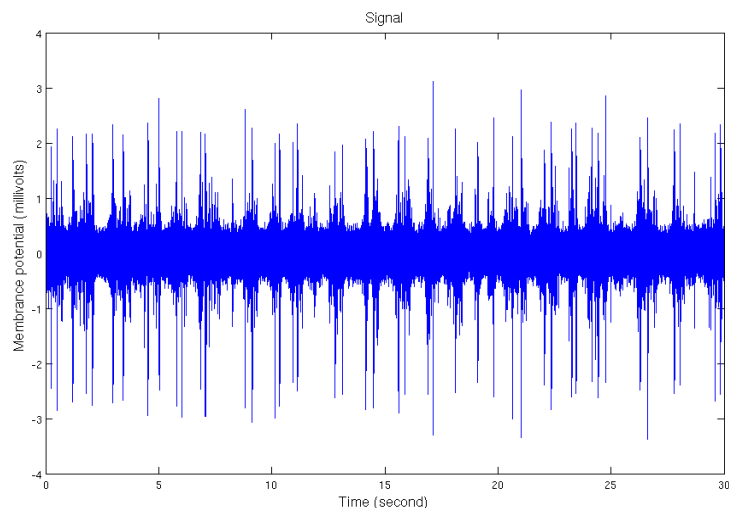


Figure 2.3: 神經訊號

研究神經細胞訊號有數種方法，其中一種常見的方式是利用微電極紀錄這些細胞的膜電位變化。Figure 2.3 是一段 30 秒的真實訊號，由微電極所紀錄下來的神經細

胞放電訊號。一般常看到的動作電位是 spikes 訊號，我們由 Figure 2.3 當中選取一個 spike 訊號於 Figure 2.4。對照 spike 訊號以及上述提及的絕對不反應期與相對不反應期，在沒有額外刺激之下，1 ~ 4 毫秒內的確沒有其它的動作電位發生。

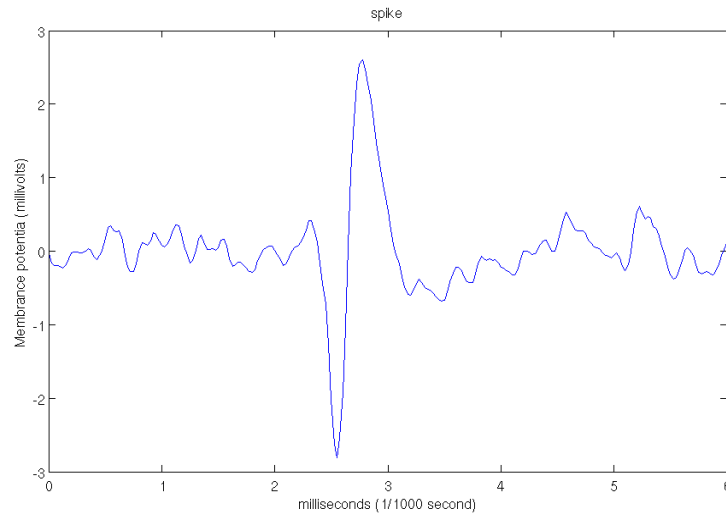
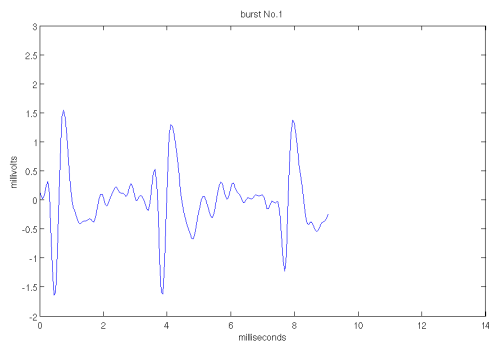


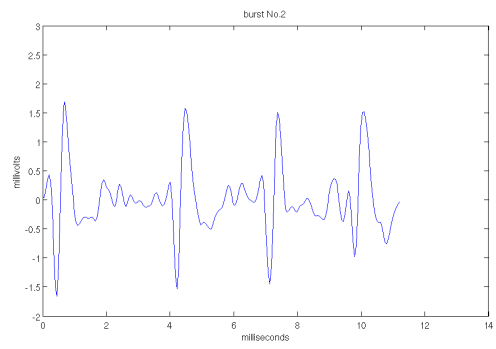
Figure 2.4: 發生於 8.814~8.82 秒之間的棘波 (spike) 訊號

腦部外層有一層大腦皮質 (Cerebral Cortex)，在此處有大量的神經細胞聚集。在大腦皮質中紀錄到的神經訊號多數為 spikes，當微電極再往大腦內部探測時，紀錄到了另外一種與 spikes 不同的訊號，如 Figure 1.2，這種形式的訊號稱之為 bursts。為了方便起見，之後的討論將一次膜電位變化視為一個 unit，Figure 1.2 就是一組由 3 個 units 所組成的 burst 訊號。由 Figure 1.2 觀察 bursts 訊號可以看到兩個明顯的特徵：第一、bursts 會接連發生數個 units。第二、units 振幅會逐次衰減。本文所使用的 bursts 訊號是國立宜蘭大學生物機電工程學系的蔡孟利教授所提供，由於 bursts 的特徵尚未量化，因此要找出 bursts 需要透過直接觀察。感謝蔡教授在眾多訊號中幫忙篩選出 112 個具有明顯特徵的 bursts 訊號。

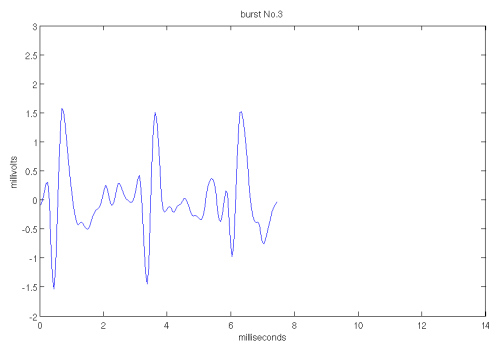
為了對上述提到的兩個 bursts 特徵更有印象，在 Figure 2.5 中我們列舉幾個 bursts 訊號。在本文接下來的章節之中，將分析蔡教授提供的 112 個 bursts 訊號，找出 bursts 的特徵並給予定性定量的描述。



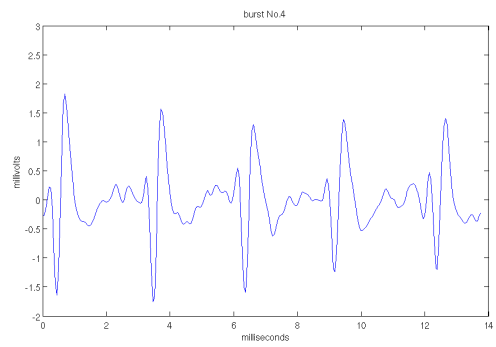
(a) Burst No.1



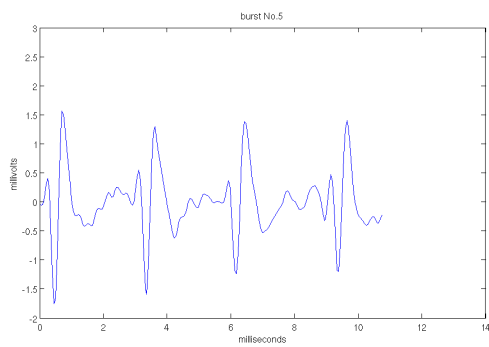
(b) Burst No.2



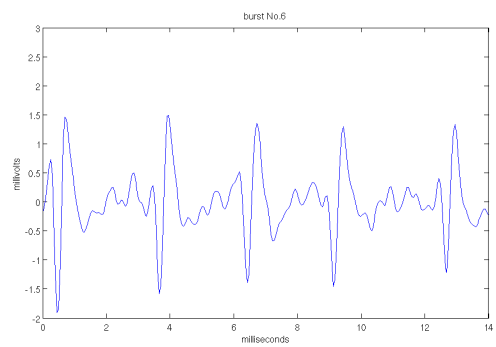
(c) Burst No.3



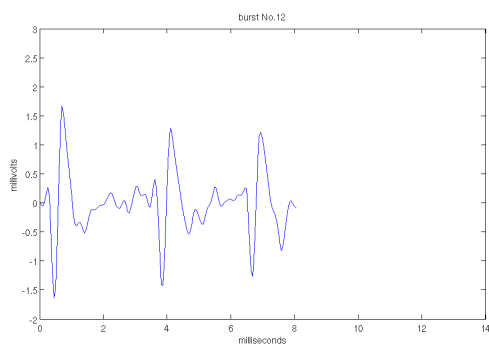
(d) Burst No.4



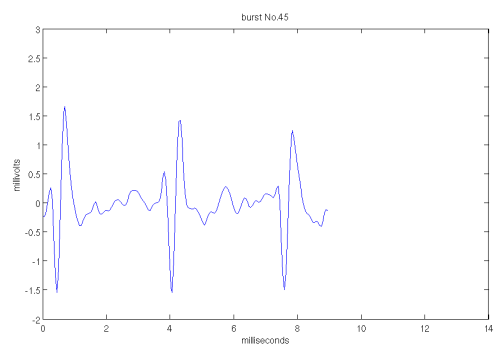
(e) Burst No.5



(f) Burst No.6



(g) Burst No.12



(h) Burst No.45

Figure 2.5: Bursts 訊號

第 3 章 實驗訊號處理與分析

本章要說明詳細的處理過程。由於 bursts 是尚未有明確定義的訊號種類，因此並沒有辦法以電腦製造大量的模擬訊號，在本文中所處理的資料都是實驗中所紀錄下來的真實訊號。

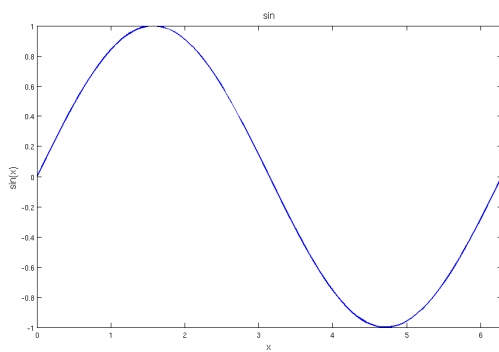
在此處先簡單介紹接下來將說明的內容。3.1 節將訊號降低取樣頻率 (down sampling) 以及去雜訊 (denoising) 的前置處理。3.2 節從蔡孟利教授所挑選出來的 bursts 訊號中，分析並尋找 bursts 的特徵並量化之。找到特徵並設定好篩選 bursts 的條件後，我們在 3.3 節使用這些條件再次檢驗上述已知為 bursts 的訊號，看看設定的條件是否有效。3.4 節將 3.1 節的前置處理，以及 3.2 節的篩選條件，實際套用到原始訊號中做一次完整的訊號處理。

3.1 訊號的前置處理

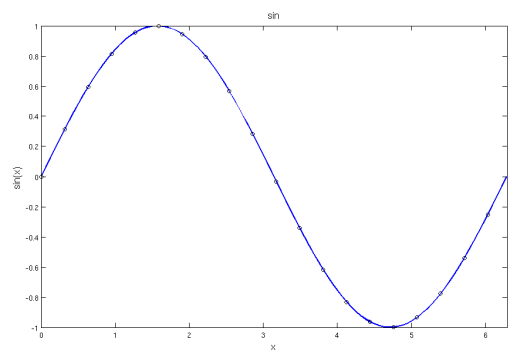
從第 2 章可以知道，取得神經訊號的方法，是利用探針紀錄在實驗過程中細胞內外的膜電位變化。當取樣頻率設定為 100KHz (100,000 Hz) 時，所代表的意思是每 1/100000 秒就紀錄一次電位。以 30 分鐘的紀錄來說，在 100KHz 的取樣頻率下，會有 1.8×10^8 個數值，如果每個數值都以雙精度浮點數來儲存，那麼 30 分鐘的紀錄檔將高達 1GB (Gigabyte)，這是個相當驚人的資料量。雖然高的取樣頻率能夠紀錄較多細節，但同時也增加了資料處理的難度，因此我們需要降低取樣頻率 (down sampling)，本文以下簡稱為「降維」。

舉個例子，現在有一個以 100 個取樣點繪製的 sin 圖形，如 Figure 3.1 (a)。如果要將取樣點縮小至 20 點，也就是說，在原本的 100 個點中只需要選取 20 個點做為代表，那麼最直接的方式是每隔 5 點取 1 點，如 Figure 3.1 (b)。有了新的取樣點之後，將每個點再重新連接起來做為新的圖形，如 Figure 3.1 (c)。Down sampling 會遺失部份細節的資訊，如 Figure 3.1 (d)。雖然 down sampling 會讓資料不完整，但可以大幅降低資料量，縮短處理的時間，因此，如果高取樣頻率紀錄到的細節資訊不影響資料分析的話，那麼 down sampling 是好的前置處理。

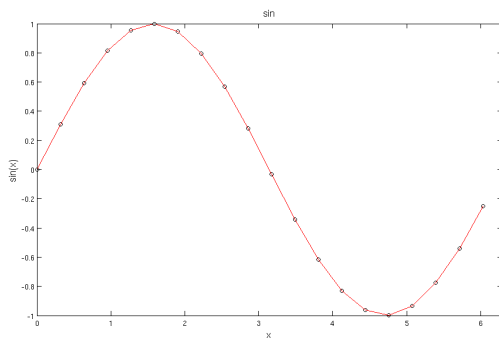
在捨去細節而不影響主要特徵的情況之下，經過適當 down sampling 降低資料量後，接下來的處理便是去雜訊 (denoising)。只要紀錄到的訊號資料，不是研究或



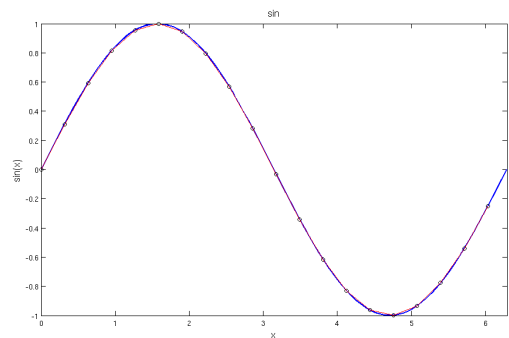
(a) 以 100 個取樣點



(b) 降成 20 個取樣點



(c) 以新的取樣點重新繪製圖形



(d) 100 個取樣點與 20 個取樣點比較

Figure 3.1: 將 $\sin(x)$ 由 100 個取樣點 down sampling 至 20 個取樣點

分析的目標，都可以叫做雜訊。一般常用的去雜訊方法不外乎低通濾波器（low-pass filter）以及高通濾波器（high-pass filter）。

低通濾波器（low-pass filter），顧名思義低的頻率可以通過這一個的過濾器，也就是說將會刪除高的頻率。換言之，高通濾波器（high-pass filter）將過濾掉低的頻率。Low-pass filter 以及 high-pass filter 只是單純地過濾高頻或低頻，但是雜訊可能同時存在於任何的頻率上，因此接下來我們將使用 shrink 的方式來 denoise。

在說明 shrink 濾波之前，先簡單介紹所謂的頻譜圖（spectrum）。傅立葉轉換 [3] 是一個從時間域轉換到頻率域的運算。訊號透過傅立葉轉換後所得到的數值取絕對值，其代表了各頻率的能量，將這些數值表示於一張縱軸為能量強度，橫軸為頻率的圖上，這就是頻譜圖（spectrum）。以白雜訊（White noise）為例，它是一種在時間域為常態分佈（Normal distribution），頻率域為均勻分佈（Uniform distribution）的訊號。Figure 3.2 (a) 是 10 秒鐘 20KHz 的白雜訊在時間域中的表現，Figure 3.2 (b) 是經過傅立葉轉換後在頻率域中的表現。

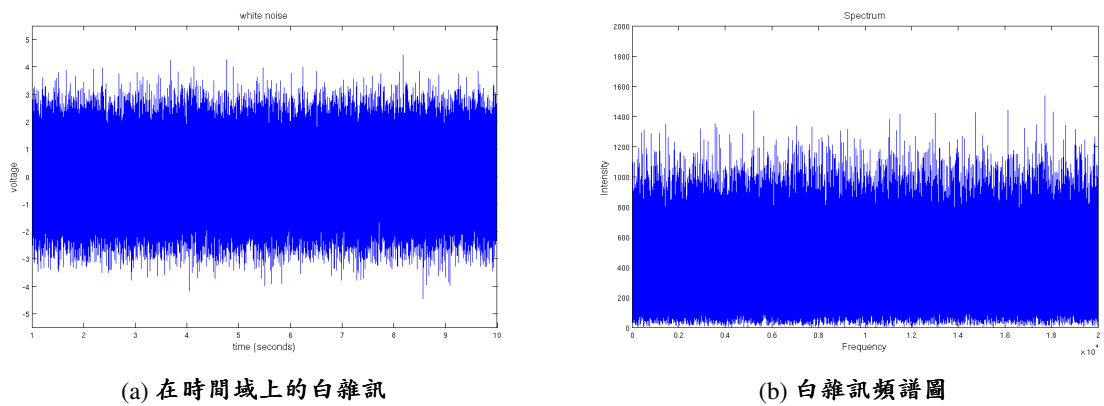


Figure 3.2: 白雜訊模擬圖

Shrink 濾波 [3, p.268] 是眾多 denoising 方式中的其中一種，與 low-pass filter 以及 high-pass filter 不同的地方在於，shrink 考量到在任何的頻率上都可能有雜訊存在，因此應該要同時過濾了高頻和低頻。在 Figure 3.2 (b) 可以看到，白雜訊的頻率並不是集中在高頻或者低頻，它是呈現均勻分佈於每個頻率。神經訊號的雜訊並不是白雜訊，它在時間域的直方圖會在正規分佈以外於 0 附近有異常的突起，稱此為中央突起 [2, p.16]，如 Figure 3.3。

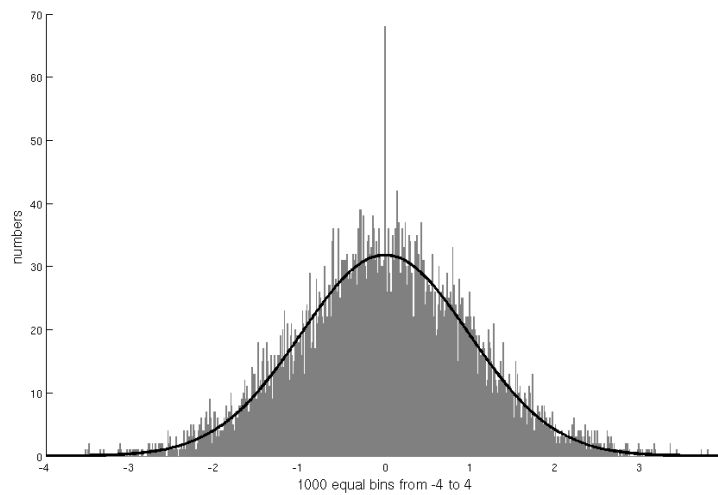
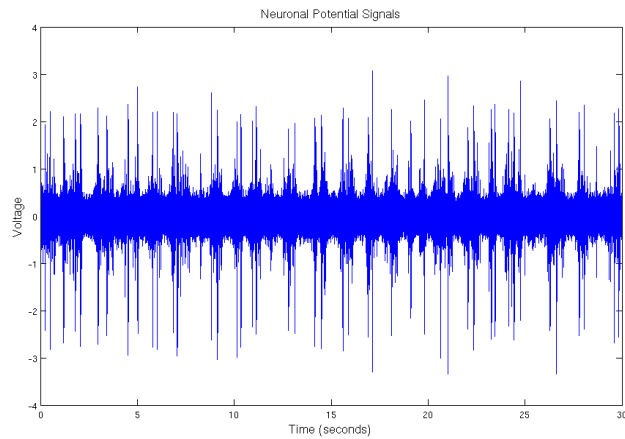


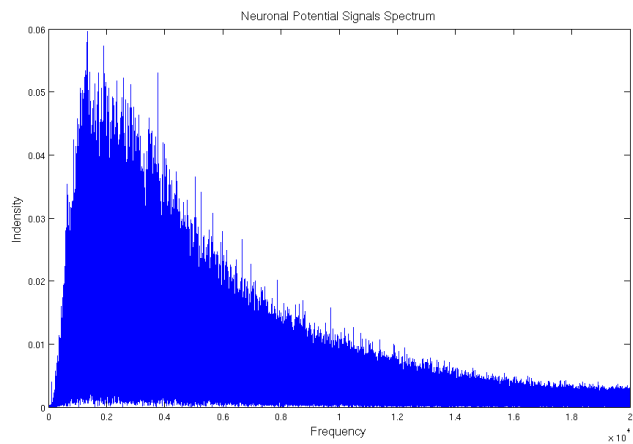
Figure 3.3: 神經訊號雜訊組成示意圖

雖然神經訊號的雜訊不完全符合白雜訊的模形，但在尚未瞭解雜訊組成的情況之下，若暫時忽略中央突起，是可以將其視為白雜訊以便做最初步的過濾。

本章節使用的數據，是由國立宜蘭大學生物機電工程學系的蔡孟利教授所提供



(a) 截取 30 秒訊號



(b) 訊號頻譜圖

Figure 3.4: 實驗中紀錄的神經元電位訊號圖

的一段長時間的神經細胞放電紀錄，為了方便後續的說明，我們任選一段訊號做為代表。

Figure 3.4 (a) 是 20KHz 的 30 秒訊號，Figure 3.4 (b) 為其頻譜圖。頻譜圖中顯示，訊號主要的特徵集中在低頻的地方。從「白雜訊在頻率域中是均勻分佈」的這個定義來看，雜訊應該存在於各頻率。因此，可以透過傅立葉轉換得到頻譜圖，設定過濾閾值，再依閾值減弱各頻率的強度後，透過傅立葉逆轉換就可以得到過濾後的新訊號，此過程即為 shrink 濾波。

Figure 3.5 是真實訊號的 shrink 濾波過程，比較 Figure 3.5 (a) 和 Figure 3.5 (d)，從圖中可以看到 shrink 之後仍然保留著訊號的主要特徵，而且曲線變的更平滑（參考 Figure 3.5 (a) 和 (d) 畫框內的訊號）。雖然振幅略為變小，但並不影響判斷。

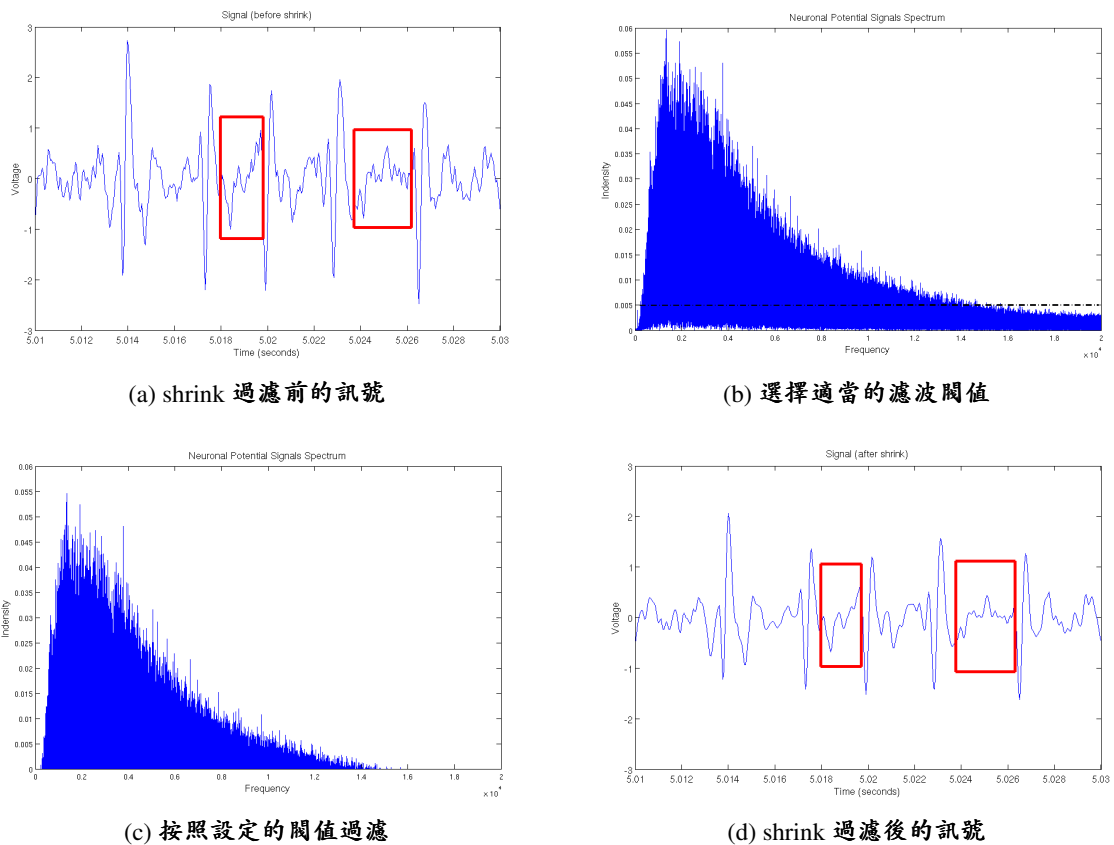


Figure 3.5: 實驗中紀錄的神經元電位訊號使用 shrink 濾波去雜訊的過程圖

3.2 分析訊號與建立篩選條件

從第一章知道，本文所要討論的主題—bursts，是一種在短時間內會接連發生數個反應的訊號。為了瞭解 bursts 的特徵，並且給每一個特徵一個清楚的定義，我們需要先以人工挑選的方式找出大量的 bursts 訊號，分析這些資料來尋找共同的特徵。國立宜蘭大學蔡孟利教授提供了 112 個來自大腦視丘且具有明顯 bursts 特徵的訊號，以下將利用這些訊號來定義三個可量化的篩選條件。

3.2.1 間隔條件

在前一章我們有提到，bursts 是一種由數個 units 所組成的訊號。從 Figure 2.5 這幾個已知為 bursts 的訊號來看，組成 bursts 的 units 個數並不固定，可能是三個，也可能是四個... 等。Figure 3.6 是一個由三個 units 所組成的 bursts，為了方便起見，我們定義 bursts 第一個發生的動作電位叫做第一個 unit (1st. unit)，第二個發生的動作

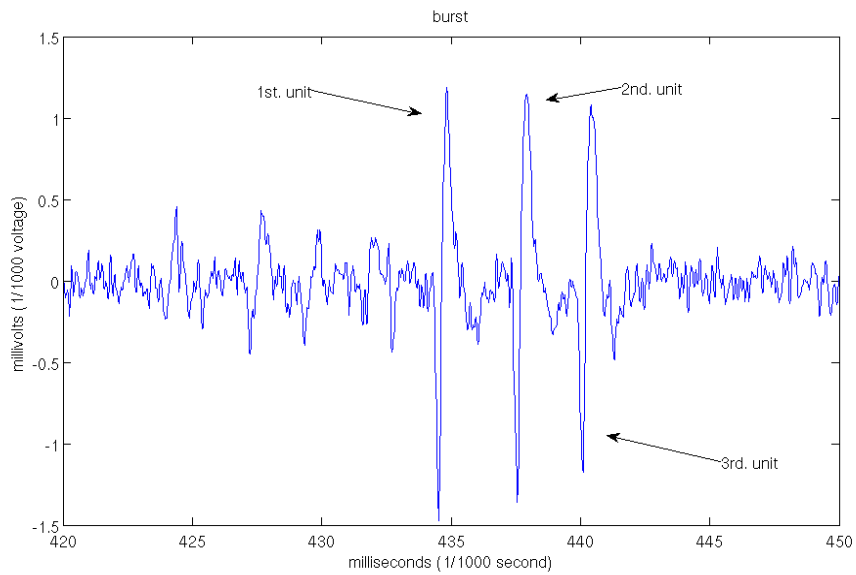


Figure 3.6: bursts 訊號圖

電位叫做第二個 unit (2nd. unit)，以此類推。從圖中可以看到，unit 與 unit 之間並不是接連，而是有時間間隔的。設定 units 的波谷做為 units 的發生時間，將每一個 burst 的所有 units 時間間隔平均，做為該 burst 的 units 的平均間隔時間。

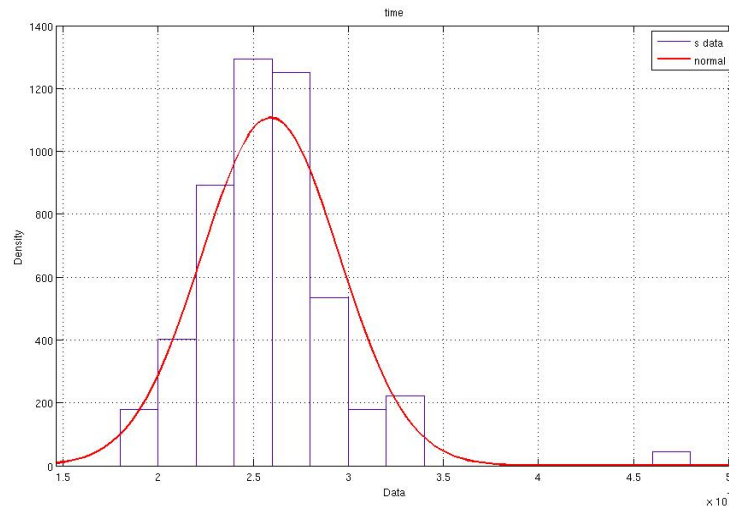


Figure 3.7: bursts 平均間隔時間直方圖

Figure 3.7 是 bursts 的平均間隔時間的統計直方圖，從圖中所示，統計結果約呈現常態分佈，且所有時間幾乎都介於 1 ~ 4 毫秒之間。參考第二章神經細胞的基本特性，以及數據統計的結果，我們將設定第一個篩選條件：「接連發生的 units，彼此的時間間隔需介於 1 ~ 4 毫秒」。

3.2.2 遞減條件

Bursts 另一個顯著的表徵是，每個 unit 的振幅會逐次遞減。令 units 的振幅大小為該 unit 的波峰與波谷的距離，以橫軸為時間，縱軸為振幅大小，每個 unit 以一個點做為表示，該點的橫軸坐標為 unit 發生的時間點，縱軸坐標為 unit 的振幅大小。

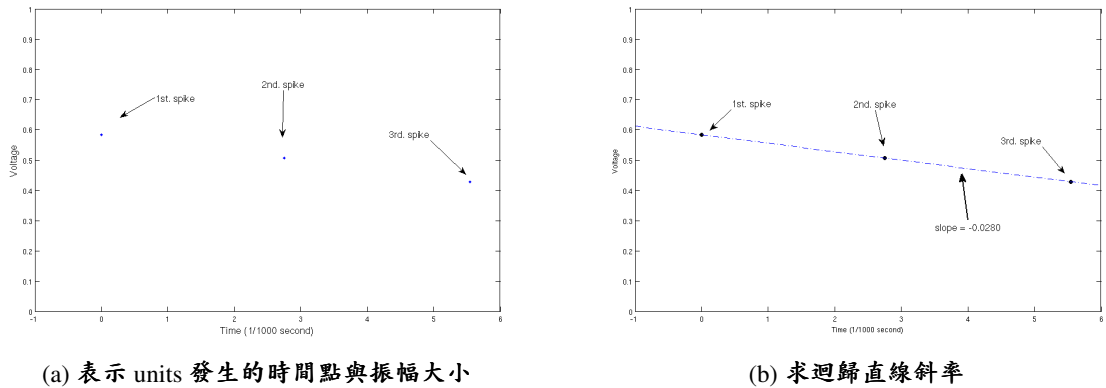


Figure 3.8: bursts 振幅遞減示意圖

按照這個方式將 bursts 的振幅變化以圖來表示，Figure 3.8 是上述方式的示意圖。Figure 3.8 (a) 裡每個點都代表一個 unit 的振幅。從圖中可以明顯看到振幅遞減的趨勢，為了更明確的說明這一點，我們使用最小平方方法在這些點中找一條迴歸直線，並計算直線的斜率。Figure 3.8 (b) 的迴歸直線斜率為 -0.0280 ，很明顯表示振幅是衰減的。現在知道了 bursts 的每個 units 振幅，以及其迴歸直線斜率。

接著，繪製 bursts 的斜率與第一個 unit 振幅的散佈圖，如 Figure 3.9。從圖中可以看到，振幅與斜率是有關係的，Figure 3.9 的相關係數是 -0.7172 ，考慮到雜訊對訊號的影響，因此這樣子的數值大小已經顯示出彼此間存在著某種關係。已知 units 振幅會下降，且振幅與其下降斜率有所關聯，那麼是否每一次振幅減少的程度是某種比例呢？答案是肯定的。我們認為下一個 unit 的振幅會是前一個 unit 振幅的 r 倍 ($r < 1$)，直到最後一個 unit 振幅為第一個 unit 振幅的 80% 為止，最後振幅的大小會再受到雜訊的影響。第四章將做一些數值模擬檢查振幅間可能的關係，並說明為何我們認為 units 振幅衰減的關係，為上面描述的簡單倍數加上雜訊影響。

組成 bursts 的 units 個數不盡相同，在已知的 bursts 當中，至少都有三個 units，令第一個 units 的振幅為 h_1 ，第二個 units 的振幅為 h_2 ，... 以此類推，最後一個 units 的振幅為 h_{last} 。檢查前三個 units 振幅，再計算其平均值與標準差，如 Table 3.1。從

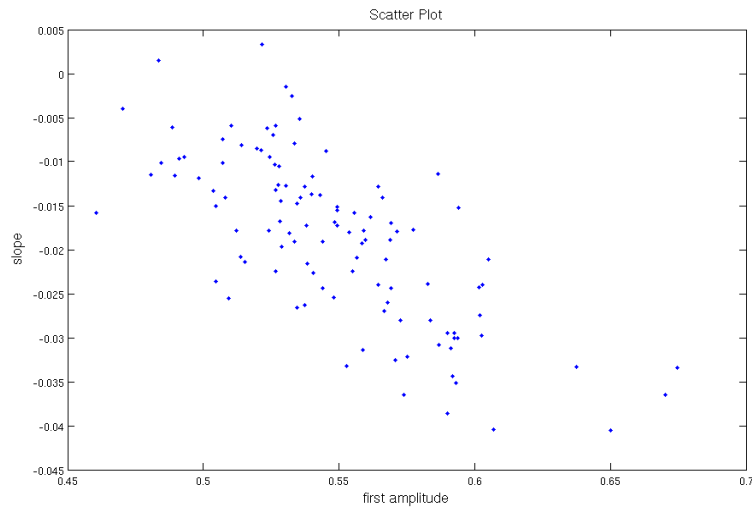


Figure 3.9: 迴歸直線斜率與第一個 unit 振幅散佈圖

	平均值	標準差
$\frac{h_2}{h_1}$	0.9076	0.0627
$\frac{h_3}{h_2}$	0.9018	0.0819
$\frac{h_3}{h_1}$	0.8179	0.0739

Table 3.1: 前三個 units 的振幅比率表

表中可以看到相鄰的 units 振幅幾乎都以 0.9 的比率下降。

在已被確認為 bursts 的訊號中，以三個或四個 units 組成 bursts 的個數較多，Table 3.2 和 Table 3.3 分別將這兩種 bursts 分開計算。在上述這些表格中可以看到一個現象，相鄰的 units 在一開始幾乎都以 0.9 的幅度逐漸衰減，直到振幅降至第一個振幅的 0.8 倍後就不再有下一個反應。所以在 Table 3.3 看到，第四個振幅和第三個振幅相當接近，其衰減的幅度不再像前幾個 units 那樣明顯。

	平均值	標準差
$\frac{h_2}{h_1}$	0.9052	0.0624
$\frac{h_3}{h_2}$	0.9042	0.0790
$\frac{h_3}{h_1}$	0.8177	0.0854

Table 3.2: 由三個 units 組成 burst 的 units 振幅比率表

	平均值	標準差
$\frac{h_2}{h_1}$	0.9187	0.0613
$\frac{h_3}{h_2}$	0.9016	0.0595
$\frac{h_4}{h_3}$	0.9482	0.0755
$\frac{h_4}{h_1}$	0.7840	0.0835

Table 3.3: 由四個 units 組成 burst 的 units 振幅比率表

當 units 振幅逐次衰減，假設 $h_{\text{last}} = 0.8h_1$ 時，在忽略雜訊的影響之下，其下降斜率為

$$\begin{aligned}
 \text{斜率} &= \frac{h_{\text{last}} - h_1}{t} \\
 &= \frac{0.8h_1 - h_1}{t} \\
 &= (0.8 - 1) \frac{h_1}{t} \\
 &= -0.2 \frac{h_1}{t}
 \end{aligned} \tag{3.1}$$

令 t 為時間時，從 (3.1) 式可以看到，當 h_1 愈大時，則斜率愈小； h_1 愈小時，則斜率愈大，兩者為負相關，正好符合 Figure 3.9 的結果。自 Figure 3.9 以及 Table 3.1~3.3，我們將設定第二個篩選條件：「units 的振幅需以 0.9 的比率逐次下降，當振幅衰減至第一個振幅的 80% 或更低時，將不再有下一個 unit 發生」。

3.2.3 形似條件

從第二章知道，每一個神經細胞產生動作電位的波形大小都不會完全相同。Bursts 是一種由同一個神經細胞所發生的訊號，雖然每個 units 的振幅會有所遞減，但是經過正規化後，其波形應該要相似。

正規化的目的是，讓每個東西都能在相同的基準點上做比較。此處所謂的「正規化」，是將每個 burst 中的 units 振幅，以第一個 unit 為基準乘上一常數，使得振幅大小相同。把組成 bursts 的 units 一一取出，經過正規化後，計算與第一個 unit 波形的相對誤差。令第一個 unit 為 s_1 ，第二個 unit 為 s_2 ，... 以此類推，則 $i = 2 \dots n$

$$\text{相對誤差} = \frac{\sqrt{\sum (s_i - s_1)^2}}{\sqrt{\sum s_1^2}} \times 100\% \tag{3.2}$$

我們將每一個已知的 bursts 訊號都以 (3.2) 式的方式做誤差檢驗，並統計所有的誤差。

在 Figure 3.11 中可以看到，幾乎所有的相對誤差都是小於 50%，由此可以驗證了同一個神經細胞所產生的訊號，其波形是相同的。依照對神經細胞的基本認識，以及數據上的顯示，因此我們將設定第三個篩選條件：「每個 units 與該 burst 的第一個

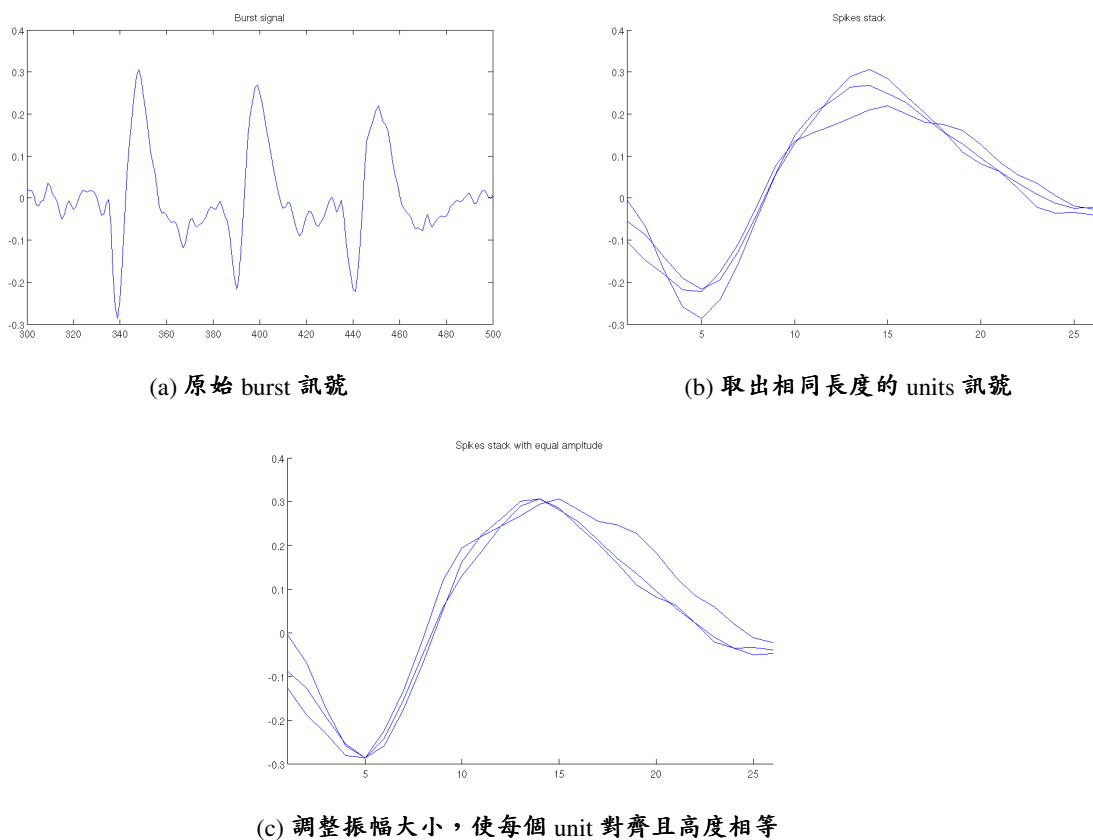


Figure 3.10: 計算組成 bursts 的 units 波形相對誤差過程圖

unit 相比，其相對誤差不可大於 50%」。

3.3 篩選條件的測試

在前一節中，我們建立了三個篩選 bursts 訊號的條件

- 接連發生的 units，彼此的時間間隔需介於 1 ~ 4 毫秒。
- units 的振幅需約以 90% 的程度遞減，直到振幅衰減至第一個振幅的 80% 或更低。
- 每個 units 與該 burst 的第一個 unit 相比，其相對誤差不可大於 50%。

在本節當中，我們將使用這三個條件來篩選 112 個已知為 bursts 的訊號，檢查最後的結果是否有誤刪。

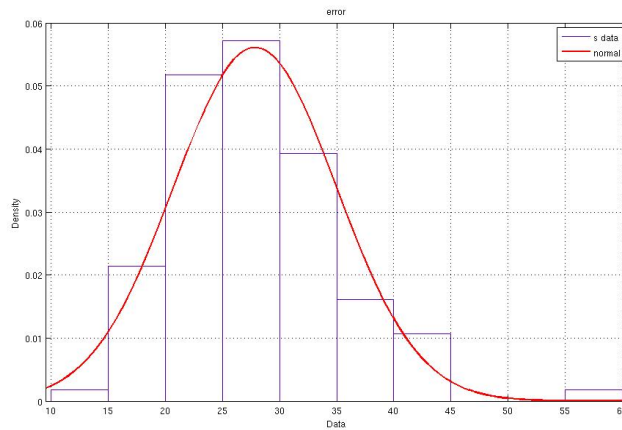


Figure 3.11: Bursts 的相對誤差統計直方圖

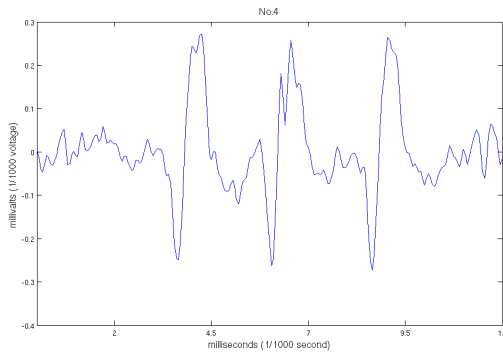
「units 的振幅會逐次遞減」，這是 bursts 的重要表徵之一。以 Figure 3.8 的方式來說明，即迴歸直線的斜率應小於 0。在測試篩選條件之前，我們利用迴歸直線的方式再次確認 112 個 bursts 振幅遞減的特徵。設定間隔條件為 1 ~ 4 毫秒、形似條件為 50%，遞減條件改以迴歸直線斜率小於 0 取代。這樣子的設定比起前一節的三個篩選條件還要來的寬鬆，此處是確認振幅遞減，因此無嚴格要求遞減幅度。

為了方便起見，我們將 112 個 bursts 編號。依上述設定檢查後，有三個 bursts 的斜率為正，如 Figure 3.12，斜率為正代表了 units 的振幅並沒有明顯遞減。在真實訊號中可能有雜訊影響，或是訊號重疊... 等因素破壞了原本訊號的特徵。我們詢問蔡孟利教授的意見，這三個訊號有可能是受到雜訊影響，導致遞減特徵不明顯甚至微幅上升，同意這三個訊號是應該被刪除的。

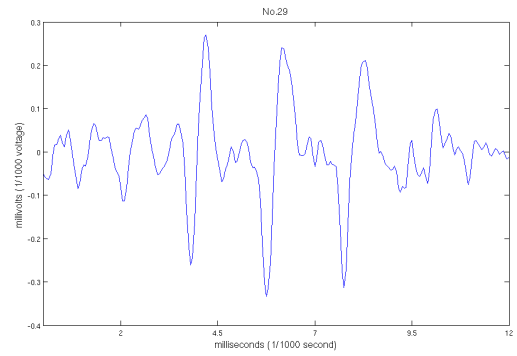
除了編號 4、29 以及 93 號 burst 以外，其餘 109 個 bursts 都有振幅遞減的特徵。在 Table 3.1~3.3 紀錄了 units 間振幅的比率平均值與標準差。我們以 Table 3.1 為例，相鄰 units 振幅的平均值約為 0.9、標準差約為 0.06 與 0.08。Figure 3.13 將第一個 unit 振幅與第二個 unit 振幅的比率，以及第二個 unit 振幅與第三個 unit 振幅的比率作直方圖，從圖可以看到呈現平均值為 0.9 的常態分佈。

接著將前一節建立好的三個篩選條件做下列設定：

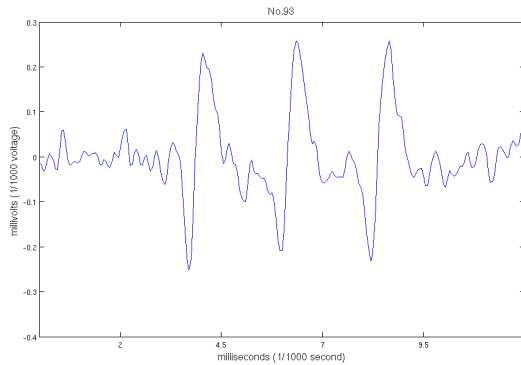
- 間隔條件：設定為 1 ~ 4 毫秒
- 遞減條件：相鄰 units 振幅遞減設定為平均值 0.9，標準差 0.07 ($0.07 = (0.06 + 0.08)/2$) 的常態分佈；最後一個 unit 與第一個 unit 振幅遞減設定為平均值 0.8，



(a) 4 號 burst



(b) 29 號 burst



(c) 93 號 burst

Figure 3.12: 112 個 bursts 中振幅未遞減的 bursts 訊號

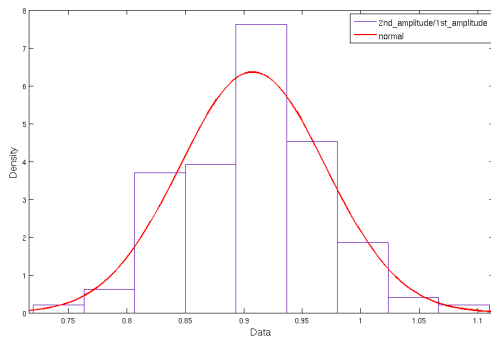
標準差 0.08 的常態分佈

- 三個標準差，即 $\frac{h_n}{h_{n-1}} \in (0.9 \pm 3 \times 0.07)$ 、 $\frac{h_{last}}{h_1} \in (0.8 \pm 3 \times 0.08)$
- 兩個標準差，即 $\frac{h_n}{h_{n-1}} \in (0.9 \pm 2 \times 0.07)$ 、 $\frac{h_{last}}{h_1} \in (0.8 \pm 2 \times 0.08)$
- 一個標準差，即 $\frac{h_n}{h_{n-1}} \in (0.9 \pm 0.07)$ 、 $\frac{h_{last}}{h_1} \in (0.8 \pm 0.08)$

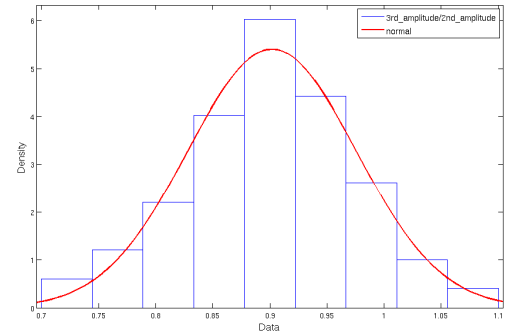
- 形似條件設定為 50%

當遞減條件設定為三個標準差時，112 個 bursts 中只刪去了編號 29 號的 bursts 訊號 (Figure 3.12 (b))。在前面已經先檢查 112 個 bursts 的振幅遞減情形，除了 29 號以外，4 號與 93 號同樣是振幅沒有遞減的訊號，而三個標準差的設定並沒有去除這兩個訊號，顯然是太寬鬆了。

當遞減條件設定為兩個標準差時，112 個 bursts 裡刪除了 11 個 bursts。11 個 bursts 除了前面的 4、29 以及 93 號之外，另外 8 個都是具有振幅遞減特徵的 bursts 訊號。



(a) 第一個與第二個 spikes 振幅比率



(b) 第三個與第二個 spikes 振幅比率

Figure 3.13: units 振幅比率直方圖

	總共個數	應通過個數	通過個數	通過率	應刪除個數	刪除個數	阻擋率
$\mu \pm \sigma$	112	109	54	49%	3	3	100%
$\mu \pm 2\sigma$	112	109	101	92%	3	3	100%
$\mu \pm 3\sigma$	112	109	109	100%	3	1	33%

Table 3.4: 篩選條件於不同設定之結果

Table 3.4 是篩選條件在不同設定之下的結果。從表中可以看到，這 112 個 bursts 幾乎都可以通過前一節所建立的篩選條件，也就是說這些條件的確可以用來偵測 bursts 訊號。在三個標準差的設定下通過率較高，不過其阻擋率低，並不能過濾掉不符合 bursts 振幅遞減特徵的訊號；在兩個標準差的設定下雖然通過率不如三個標準差高，但過篩選的效果較好。同時考量通過率與阻擋率，我們選擇以 $\mu \pm 2\sigma$ 做為較佳的遞減條件設定。

3.4 處理原始訊號以及 bursts 的分類

經過 3.1 ~ 3.3 節的前置處理以及篩選條件的建立與驗證後，在本節當中將實際來處理一筆原始訊號。

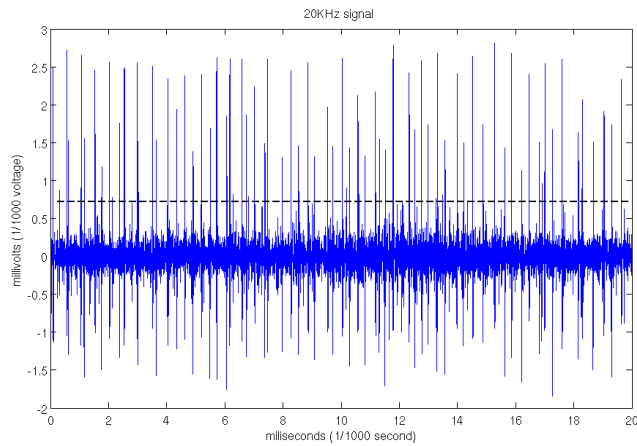


Figure 3.14: 設定 0.7 為探測動作電位的門檻值

接下來將處理的資料同樣由蔡孟利教授所提供，是一筆約莫 11 分鐘，取樣頻率為 40KHz 的原始訊號。將該筆訊號做前置處理，包括降維（從 40KHz 降至 20KHz），以及 shrink 濾波（降低 20% 的頻譜強度）。透過間隔條件後，得到許多符合該條件的訊號，凡是超過設定的電作電位閾值，且時間間隔介於 1 ~ 4 毫秒，皆符合第一個篩選條件。我們使用門檻值探測法來篩選動作電位，選擇一個超過雜訊標準的值做為門檻值，以這一筆訊號來說，門檻值設定為 0.7，如 Figure 3.14。如果恰好有兩個神經元在短時間內接連發生反應，或是較大的雜訊，都有可能符合該條件。

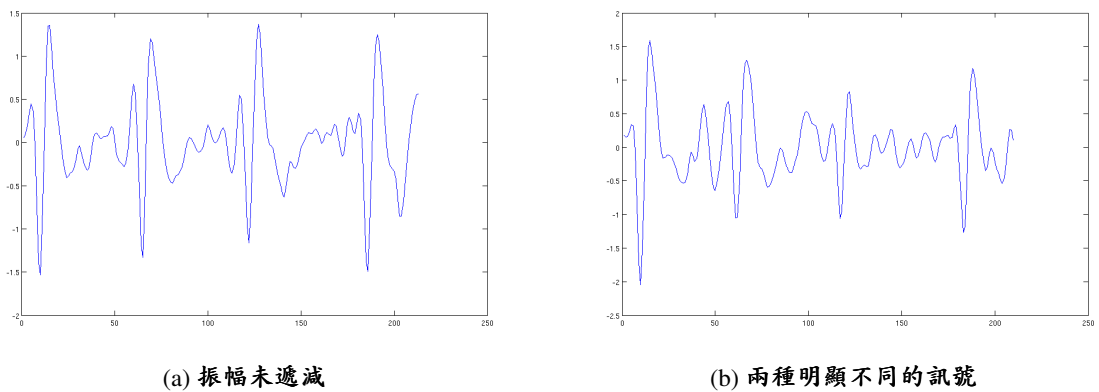


Figure 3.15: 通過第一個篩選條件的錯誤訊號

Figure 3.15 是兩個符合間隔條件，但不是 bursts 的訊號，因此，我們需要利用遞減條件。將通過間隔條件的訊號，以遞減條件加以判斷。考慮到雜訊的影響，遞減條件的比率並不是絕對的，而是約略為 90%。在前一小節說明了，平均值正負兩個

標準差的範圍內具有較平衡的通過率與阻擋率。參考前一小節中的結果，遞減條件將篩選出相鄰的 units 其振幅變化落在 $0.9 \pm 2 \times 0.07$ 的訊號。經過遞減條件的過濾，留下了 308 個符合間隔條件以及遞減條件的訊號。

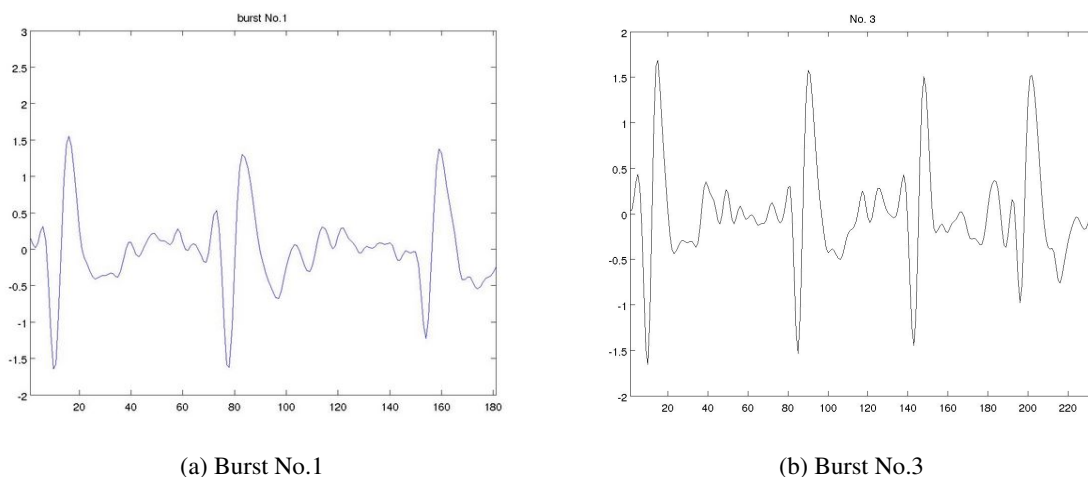


Figure 3.16: 通過間隔條件以及遞減條件的訊號

經過間隔條件以及遞減條件的過濾，已經可以找到幾乎符合 bursts 特徵的訊號，在 Figure 3.16 舉兩個做為代表。不過，仍有一種錯誤情況可能發生，如 Figure 3.17。

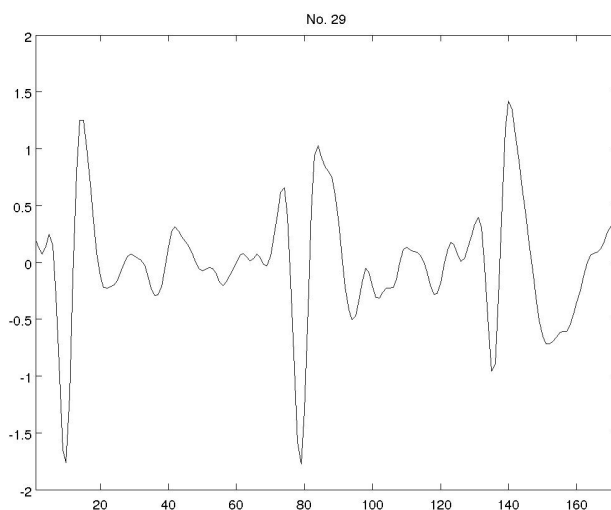


Figure 3.17: 通過間隔條件以及遞減條件的錯誤訊號

雖然 Figure 3.17 發生 units 的間隔介於 1 ~ 4 毫秒，且振幅變化也符合遞減條件，但是很明顯可以看到，第二個 unit 的波形與其它兩個有很大的不同。因此，需要再

加入形似條件。形似條件將逐一檢查 bursts 裡頭的每一個 units，確認 units 的波形是相似的，也就是說，該 burst 是由同一個神經元所產生的，如此可以排除 Figure 3.17 的情況，也就是排除不同神經元在相鄰的時間點發生反應的巧合。

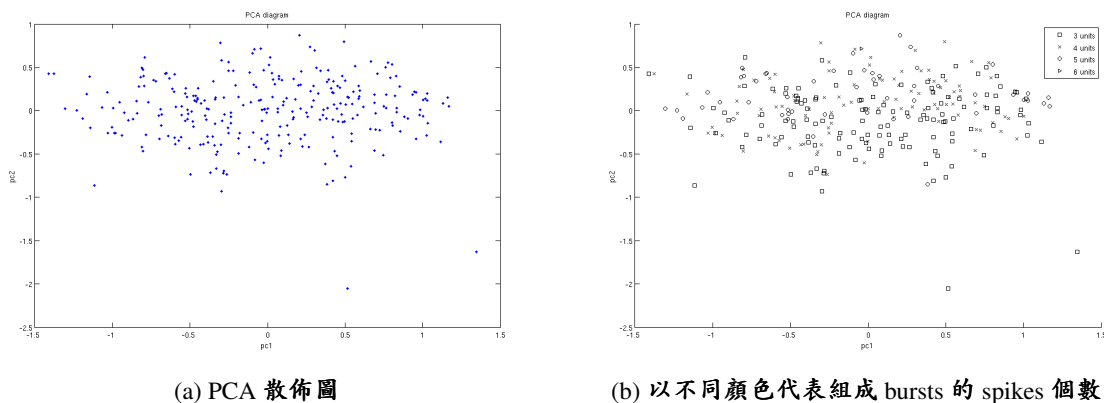


Figure 3.18: Bursts PCA 分類圖

接著將篩選出來的 283 個 bursts 訊號做分類。根據 3.2 節的形似條件的說明，我們認為 burst 是來自於同一個神經細胞，因此 burst 裡頭的 units 波形應該彼此相似，因此我們就取第一個 unit 的波形做為每個 burst 的代表波形。通常在分類 spikes 訊號（這裡指的是一般的動作電位，如 Figure 1.1 (b)）時，是使用 PCA（Principal component analysis）做分類 [4]。基於上述理由，同樣使用 PCA 為 bursts 訊號做分類，結果於 Figure 3.18。

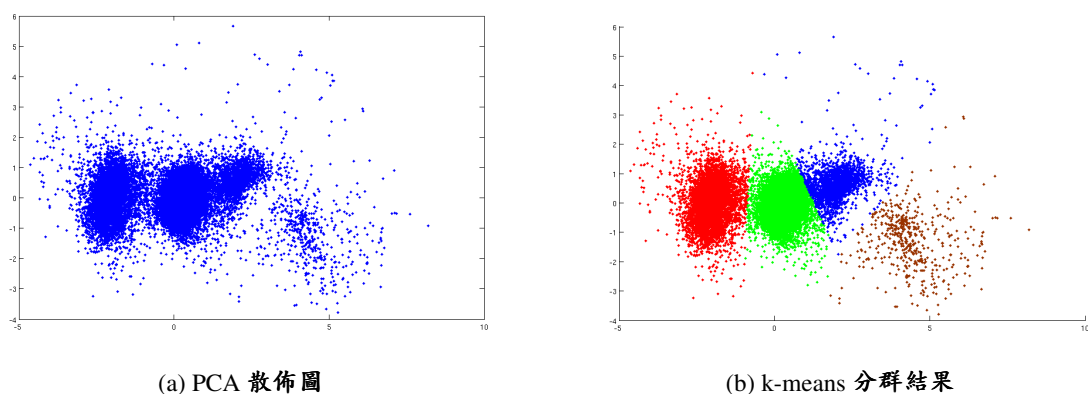


Figure 3.19: Spike (動作電位) 利用 PCA 與 k-means 分類結果圖

Bursts 的代表波形經過 PCA 計算後，從散佈圖中並沒有辦法很明確的分成一類以上。另外，一個 burst 可能由數個 units 組成，units 的個數是否可以由分佈圖中區

分出來呢？現在於分佈圖上增加 units 個數的資訊，相同 units 個數的 bursts 以同一種圖示代表。如 Figure 3.18 (b) 所示，各種數量的 units 是散落在每個地方，我們並不能透過 PCA 來區分 spikes 的個數。

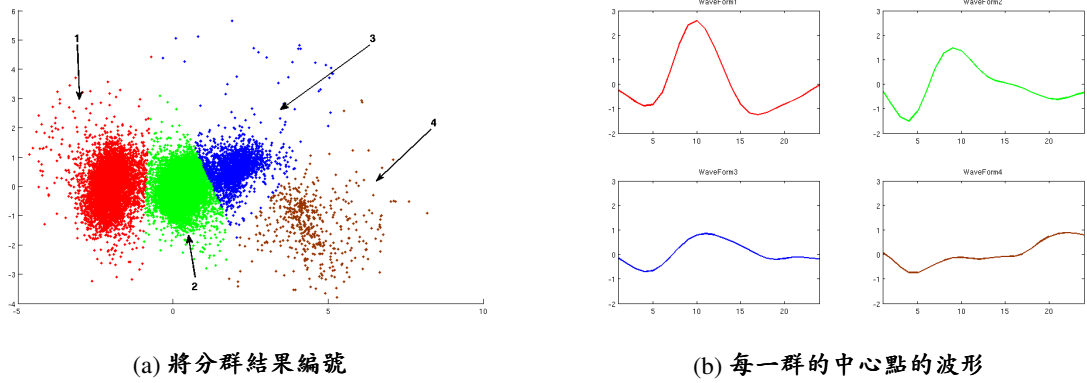


Figure 3.20: 四群動作電位的代表波形

接下來我們考察這些 bursts 是否的確由某個神經細胞所產生。首先，從同一筆訊號中找出一般的 spike 訊號並且分類。

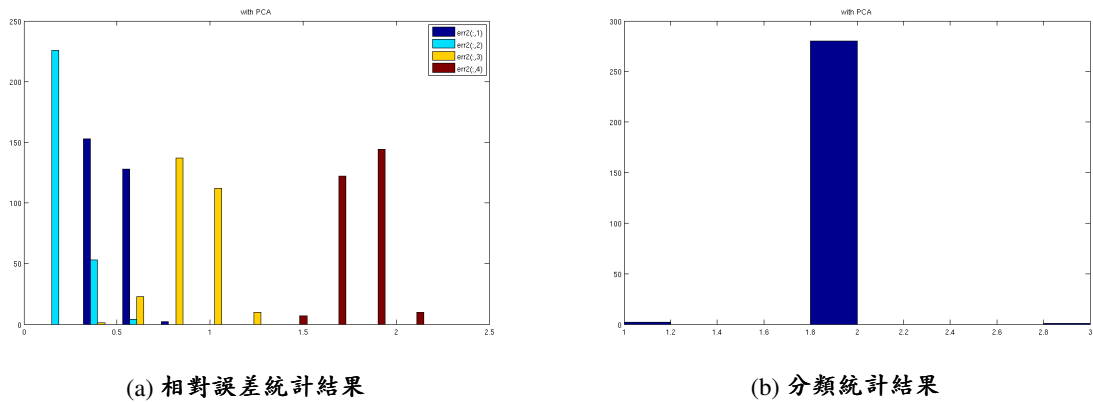


Figure 3.21: Bursts 波形與動作電位波形比對結果

Spikes 訊號透過 PCA 的運算後，選出最具有區別力的第一主成份以及第二主成份，分別置於 x 軸與 y 軸，如 Figure 3.19 (a)。得到 spikes 的散佈圖後，利用 k-means 分群法 (k-means clustering) 來做分群，將 Figure 3.19 (a) 分為四群，如 Figure 3.19 (b)。選擇每一群的中心點，以中心點的波形做為該群的代表波形。再將 bursts 的波形與動作電位分類出來的四種波形，一一比對分類後，再將分類至該動作電位的 bursts 波形，與代表波形繪製於同一張圖上比較其相似度。比對後的結果和波形如

Figure 3.20 和 Figure 3.21 。

從圖中可以看到，在這筆資料中所篩選出來的 bursts，多數都對應到某一種動作電位的波形，也就是說，某一個神經細胞除了會產生基本的 spike 訊號之外，還會產生 bursts 訊號。

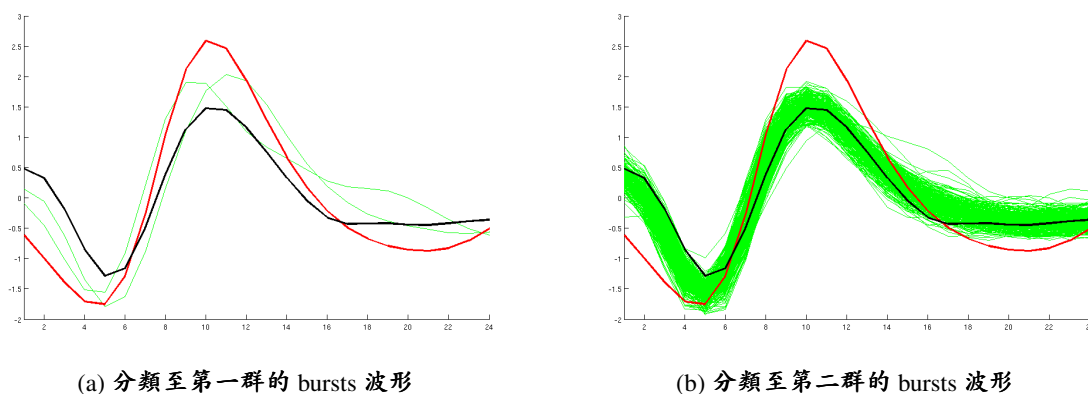


Figure 3.22: Bursts 波形與動作電位波形比較

最後的處理結果告訴我們，這一筆原始訊號紀錄到了三個動作電位的振幅超過 0.7 的神經細胞，三個神經細胞都會產生 spikes 訊號，其中有一個神經細胞不僅產生 spikes 訊號，也產生 bursts 訊號。也就是說，一個神經細胞所發生的動作電位可以有 spikes 和 bursts 兩種訊號模式。

第 4 章 振幅衰減的統計模型

Bursts 訊號的一個明顯特徵是振幅衰減，在 3.2 節中，為該特徵下了一個結論：「bursts 的每個 unit 的振幅會是前一個 unit 振幅的 r 倍 ($0 < r < 1$)，最後表現在訊號上的振幅大小會受到雜訊的影響」。本章將利用統計學上的觀念來做數值模擬 (Numerical Simulations)，以下我們先建立三種統計模型，然後，檢查三種模型所模擬的振幅衰減現象，再對照真實數據選擇最符合事實的模型，當作 bursts 訊號之振幅衰減的統計模型。

常態分佈是一種很重要的機率分佈模型，許多自然科學或是工程上的現象都符合這樣的隨機分佈模型。以下將利用常態分佈的特性做三種數值模擬，觀察資料群之間的相關性。下面簡介本章將使用到的資料模擬方式。

4.1 節模擬兩個常態分佈的隨機變數相乘之結果。4.2 節模擬平均值之比值為 r 的常態分佈隨機變數相乘之結果。4.3 節模擬一個常態分佈的隨機變數，乘以常數再加上白雜訊後的結果。4.4 節將前三節的結果對照真實數據。

在數據模擬之前，先說明一些定義與名詞。假設有兩群資料點，分別是 $X = \{x_1, x_2, \dots, x_n\}$ 、 $Y = \{y_1, y_2, \dots, y_n\}$ 。 \bar{x} 是 x_i 的平均值， \bar{y} 是 y_i 的平均值。 X 、 Y 之間的相關係數與散佈圖如下。

- 相關係數 (Correlation Coefficient)

$$\text{相關係數} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (4.1)$$

相關係數將介於 -1 與 1 之間，若所得到的相關係數的絕對值愈接近 1 ，則代表 x_i 、 y_i 之間的相依性愈大， X 與 Y 高度相關，反之，若愈接近 0 ，則低度相關。

- 散佈圖 (Scatter Plot)

將 x_i 置於橫軸 y_i 置於縱軸，則 X 、 Y 的散佈圖如 Figure 4.1。

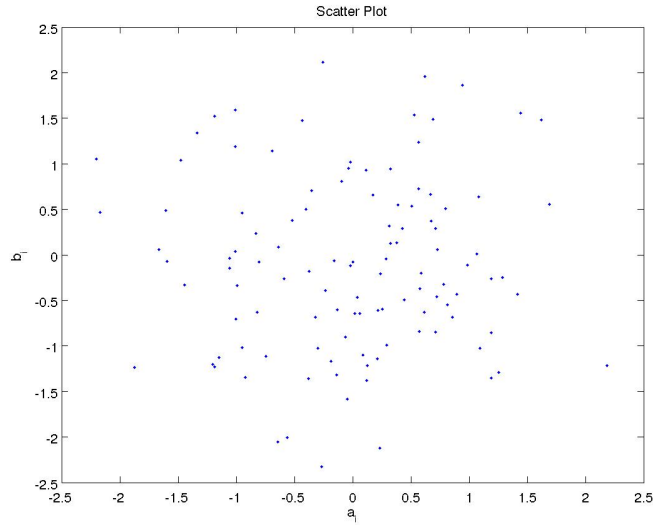


Figure 4.1: 散佈圖

爲了方便後續的說明，我們定義下列變數：

- 令 X 爲從平均值 μ_X ，標準差 σ_X 的常態分佈中，隨機取出 n 個資料點， $X = \{x_1, x_2, \dots, x_n\}$ ， x_i 是隨機變數 X 每一次取樣所得到的樣本的數值。依相同的方式分別產生 Y 、 R 、 X' 、 Y' ，這些常態分佈的平均值分別設定爲 μ_Y 、 μ_R 、 $\mu_{X'}$ 、 $\mu_{Y'}$ ，標準差分別爲 σ_Y 、 σ_R 、 $\sigma_{X'}$ 、 $\sigma_{Y'}$ 。

4.1 數值模擬一： $Y = RX$

假設 X 是一個平均值爲 μ_X 標準差爲 σ_X 的常態分佈隨機變數，代表 bursts 中第一個 units 的振幅； R 是另一個平均值爲 μ_R 標準差爲 σ_R 的常態分佈的隨機變數，代表的是相鄰振幅間的比率。令 Y 代表 bursts 中第二個 units 的振幅，而 Y 由 X 乘以 R 產生。

$$\begin{aligned}
 Y &= \{y_1, y_2, \dots, y_n\} \\
 &= R \times X \\
 &= \{x_1 r_1, x_2 r_2, \dots, x_n r_n\}
 \end{aligned} \tag{4.2}$$

從 (4.2) 可以看到， y_i 是 x_i 的 r_i 倍。因爲 x_i 是符合常態分佈的， μ_X 和 σ_X 便決定了 x_i 可能的範圍。舉個例子來說，令 $\mu_X = 1$ ， $\sigma_X = 0.05$ ，那麼 x_i 會有 68% 落在

(0.95, 1.05) 之間。如果把區間加大至兩個標準差，(0.9, 1.1) 之間將含蓋 95% 的 x_i 。再加大至三個標準差的話，那麼 (0.85, 1.15) 將包含 99% 的 x_i ，幾乎是全部了！

r_i 也是常態分佈，從上面的例子可以知道，標準差的大小決定了資料散佈的範圍。現在令 $\mu_R = 0.9, \sigma_R = 0.01$ ，那麼可以說 r_i 幾乎介於 (0.87, 0.93) 之間，這是一個很小的範圍。Matlab 內建的 randn() 函式 (Matlab functions) 可以產生符合常態分佈，且平均值為 0 標準差為 1 的隨機變數。在此處使用的 Matlab 版本是 7.6.0，randn() 函式是以 Marsaglia's Ziggurat algorithm 來產生亂數，每一次使用 randn() 函式皆不重新設定產生器的初始狀態以及種子，以避免個人對數值選擇的偏好影響實驗結果。根據這個例子的平均值與標準差，利用 randn() 函式隨機產生 x_i 以及 r_i ，接著計算 y_i 。Figure 4.2 (a) 圖是 X 、 Y 的分佈圖。 σ_R 小，也就是說 r_i 的分佈集中，在這樣子

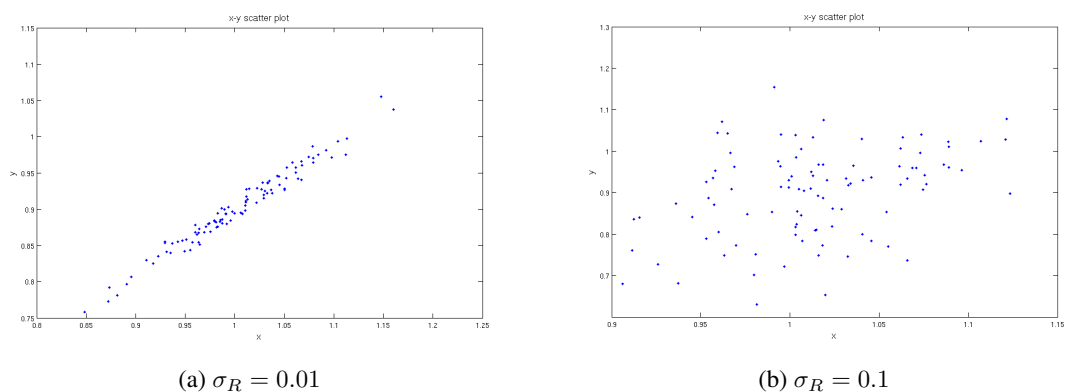


Figure 4.2: XY分佈圖

的情況之下， Y 相當於是 X 乘上一個常數。從 Figure 4.2(a) 可以看到分佈圖幾乎為直線，在這個例子中， X 、 Y 的相關係數約 0.9845。現在假設 r_i 是常數並令它為 r ，且

$r > 0$ 。按照前面相關係數的定義可以得到下面結果：

$$\begin{aligned}
 \text{相關係數} &= \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \\
 &= \frac{\sum_{i=1}^n (x_i - \bar{x})(rx_i - r\bar{x})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (rx_i - r\bar{x})^2}} \\
 &= \frac{r \sum_{i=1}^n (x_i - \bar{x})^2}{r \sum_{i=1}^n (x_i - \bar{x})^2} = 1
 \end{aligned} \tag{4.3}$$

若 $r < 0$ ，則：

$$\begin{aligned}
 \text{相關係數} &= \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \\
 &= \frac{\sum_{i=1}^n (x_i - \bar{x})(rx_i - r\bar{x})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (rx_i - r\bar{x})^2}} \\
 &= \frac{r \sum_{i=1}^n (x_i - \bar{x})^2}{-r \sum_{i=1}^n (x_i - \bar{x})^2} = -1
 \end{aligned} \tag{4.4}$$

(4.3)(4.4) 說明了當 y_i 是 x_i 的某倍數時，相關係數會是 1 或 -1 。換句話說，當相關係數的絕對值愈接近 1 表示兩者的相關性愈高；愈接近 0 代表相關性愈低。Figure 4.2 (a) x_i 與 y_i 的相關係數絕對值約 0.9845，是高度相關的。Figure 4.2 (b) 是 $\sigma_R = 0.1$ 的 XY 分佈圖，可以看到當 σ_R 較大，也就是 r_i 的分佈較為不集中時， x_i 和 y_i 的分佈情況也顯得比較離散，在 Figure 4.2 (b) 的例子中，相關係數約 0.3659。

在這裡可以得到一個結論，當 X 、 R 為獨立的常態分佈且 $Y = RX$ 的情況

下， X 和 Y 的相關係數與 R 的離散程度相關，當 R 分佈集中，也就是 σ_R 愈小時， X 相當於乘上一個常數，由 (4.3)(4.4) 可知，如此 X 、 Y 的相關係數絕對值會很接近 1；反之，若 σ_R 愈大時， X 、 Y 的相關係數絕對值會接近 0。

4.2 數值模擬二： $\mu_Y = r\mu_X$

X 同樣代表第一個 units 的振幅，與前一小節的設定相同。 Y 是一個平均值為 μ_Y 標準差為 σ_Y 的常態分佈隨機變數，代表第二個 units 的振幅。常態分佈的統計直方圖會呈現鐘型曲線，平均值即為曲線的高峰位置，標準差為曲線的寬窄程度。

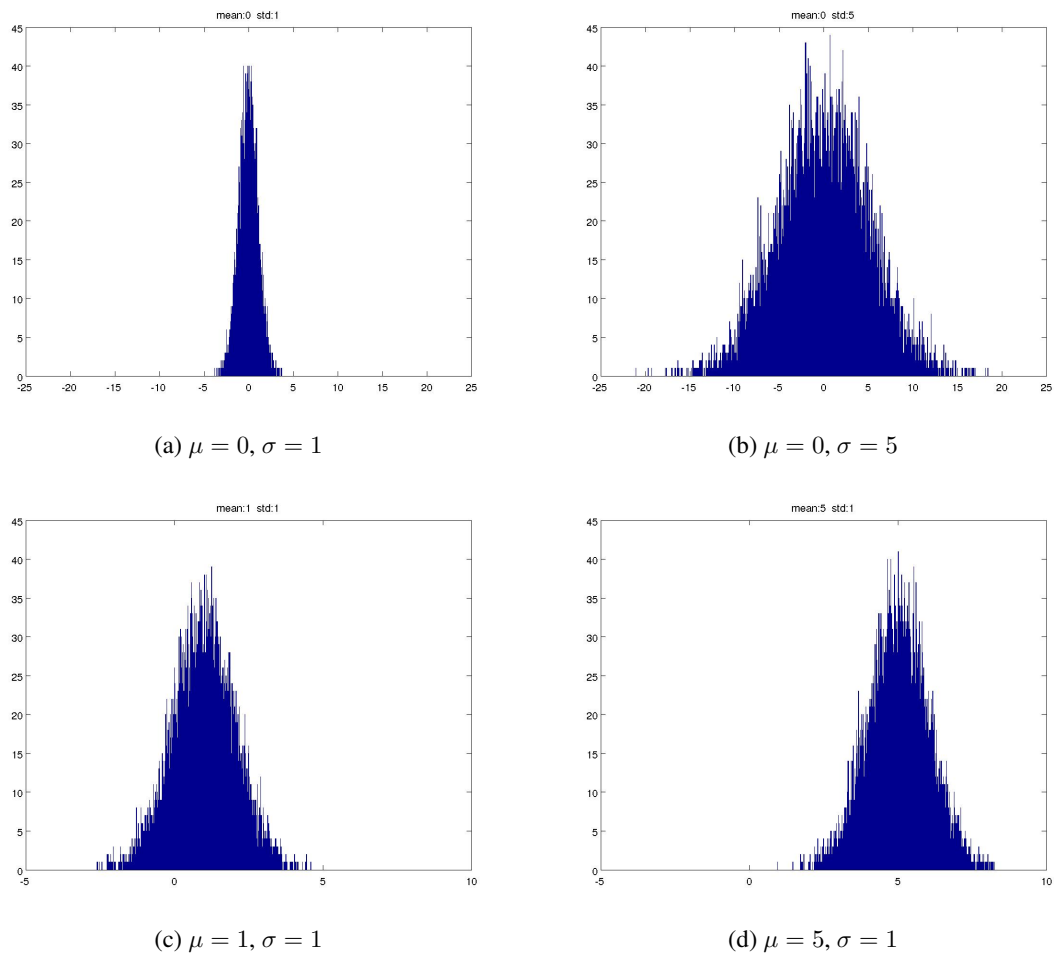


Figure 4.3: 各種常態分佈的統計直方圖

在 Figure 4.3 (a) 以及 Figure 4.3 (b) 皆固定平均值為 0，(a) 圖是標準差為 1 的統計直方圖，(b) 圖是標準差為 5 的統計直方圖，從這裡可以明顯地看到，不同的標準差

代表著不同的分佈情形。Figure 4.3 (c) 以及 Figure 4.3 (d) 固定標準差為 1，(c) 圖平均值為 1，(d) 圖平均值為 5，這兩個圖的分佈情形相同，但高峰處的發生點位置不同。

假設 X 和 Y 的分佈情況相同，也就是 $\sigma_X = \sigma_Y$ ，但平均值不同。這個想法和 Figure 4.3 的 (c), (d) 兩張圖相似。我們同樣利用 Matlab 產生樣本點 x_i 和 y_i ，令 x_i 與 y_i 具有相同的分佈，且平均值相差 r 倍， r 是常數。根據上面的設定可以寫成下列式子：

$$\mu_Y = r\mu_X \quad (4.5)$$

$$\sigma_Y = \sigma_X \quad (4.6)$$

給定 μ_X, σ_X 以及 r 後，依照 (3.5), (3.6) 可得到 μ_Y, σ_Y 的值。假設 $\mu_X = 1, \sigma_X = 1, r = 0.9$ ，Figure 4.4 是產生 x_i 與 y_i 後的分佈圖，在這個例子中相關係數約 -0.0588 。

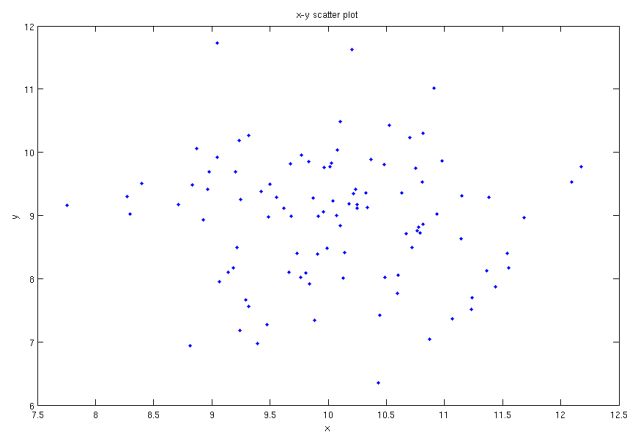


Figure 4.4: 模擬方式二的 XY 散佈圖

從圖中可以看到資料點分散不集中，因此 X 、 Y 的相關係數小，表示兩者沒有相關性。

數值模擬得到 Figure 4.4 的原因在於，雖然 x_i 與 y_i 同樣是常態分佈，且標準差相同，也就是說離散的程度相同，但是並沒有辦法保證 x_i 和 y_i 相對應的數值在鐘型曲線中與平均值的距離相等。舉個例子來說，假設有兩個標準差相同，但平均值不同的常態分佈，如 Figure 4.5。

Figure 4.5 (a) 是 x_i 的鐘型曲線，其平均值為 0，標準差為 1；(b) 圖是 y_i 的鐘型曲線，其平均值為 2，標準差為 1。如圖箭頭所示，假設 x_1 在平均值的左邊，且離

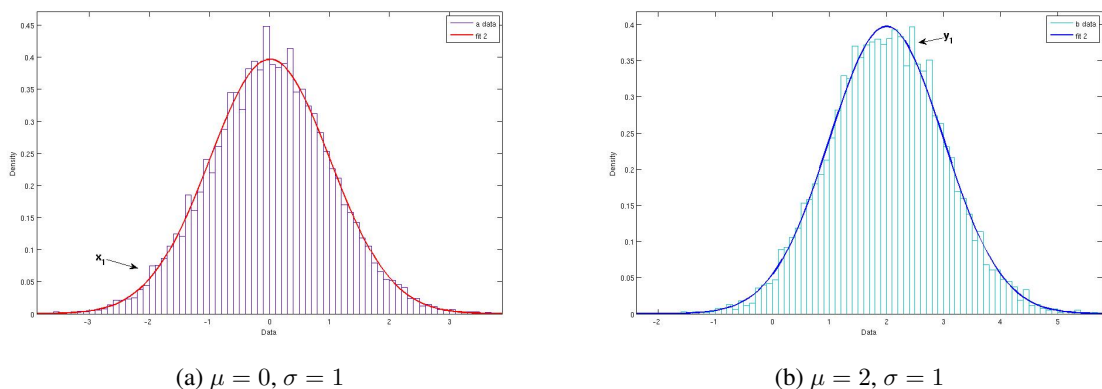


Figure 4.5: 相同分佈的常態分佈直方圖

平均值有兩個標準差之遠，而 y_1 在平均值的右邊，且很接近平均值。以這個例子來看， x_1 和 y_1 在分佈中所佔的位置相當不同。同樣的想法可以延伸到 x_i 以及 y_i 。由於 x_i 與 y_i 是隨機產生的，因此很容易發生類似 Figure 4.5 的情況，這也解釋了 Figure 4.4 的低度相關性。

在這裡可以得到一個結論，平均值相差 r 倍的兩個隨機變數，即使其分佈模型相同，甚至散離程度亦相同，但是由於兩者是獨立產生的隨機資料，並不能保證對應的資料點在分佈中落於相同的位置上，因此兩個隨機變數產生的樣本點是不相關的。

4.3 數值模擬三： $Y = rX + N$

假設 X' 是一個平均值為 $\mu_{X'}$ 標準差為 $\sigma_{X'}$ 的常態分佈隨機變數，代表第一個 units 的振幅。令 r 為常數，代表振幅下遞減的比率。 Y' 由 r 乘上 X' 所產生，代表第二個 units 的振幅。

$$\begin{aligned}
 Y' &= \{y'_1, y'_2, \dots, y'_n\} \\
 &= rX' \\
 &= \{rx'_1, rx'_2, \dots, rx'_n\}
 \end{aligned} \tag{4.7}$$

接著另外再給 X' 和 Y' 一點擾動。令 N 為平均值 μ_N 標準差 σ_N 的常態分佈隨機變

數，代表雜訊。將 N 加到 Y' ，如此得到 Y 。

$$\begin{aligned} X &= X' \\ Y &= Y' + N \end{aligned} \quad (4.8)$$

多了擾動後所產生的 x_i 和 y_i 就不再是單純的 r 倍關係。

令 $\mu_{X'} = 1, \sigma_{X'} = 0.5, r = 0.9$ ，再自 $\mu_N = 0, \sigma_N = 0.5$ 的常態分佈中產生 N 做為擾動，同樣畫出分佈圖，如 Figure 4.6。

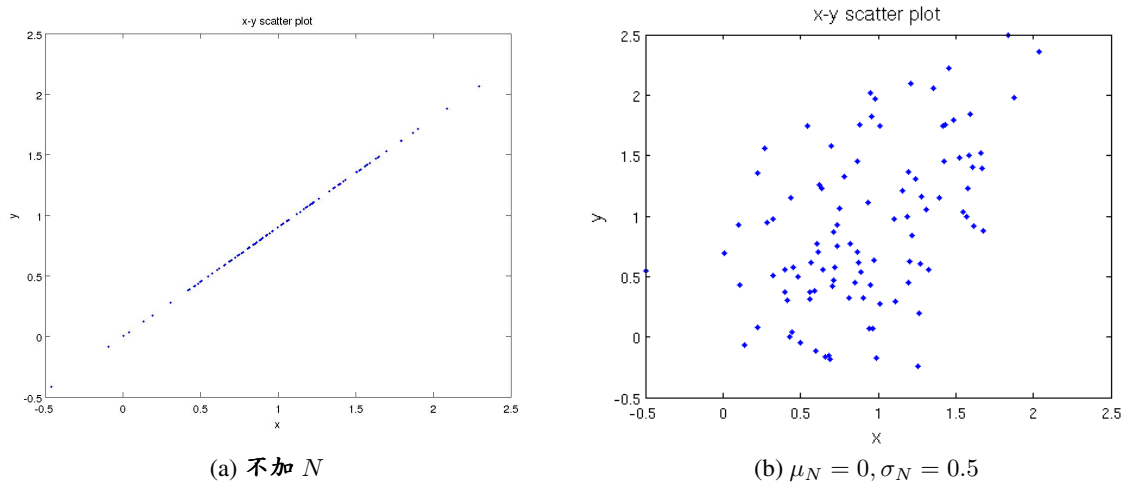


Figure 4.6: 模擬方式三的 XY 散佈圖 I

Figure 4.6 (a) 尚未加上擾動時， x_i 和 y_i 為簡單的 r 倍關係，相關係數為 1，原因在 (3.3) 式即可說明。Figure 4.6 (b) 加上擾動後，雖然資料點明顯分散許多，不過仍然可以約略呈線性分佈，相關係數 ≈ 0.4943

Figure 4.7 更改了擾動的平均值與標準差，從圖中可以看到，當雜訊的標準差愈小時，分佈圖愈集中，相關性愈高；若標準差愈大，則相關性愈低。當擾動的標準差很小，相當於加上一個常數，因此不會影像資料點之間的相對關係。反之，若標準差愈大，代表擾動的分佈範圍大，那麼資料點原始的相對關係就很容易受到擾動的影響，因此相關性就降低。

在這裡可以得到一個結論，即使第二個 units 振幅明顯地與第一個 units 振幅有關係，但雜訊所涵蓋的範圍大小將影響兩者間既有的相關性。當雜訊愈大時， X 、 Y 的相關係數愈低；雜訊愈小時， X 、 Y 的相關係數愈高。

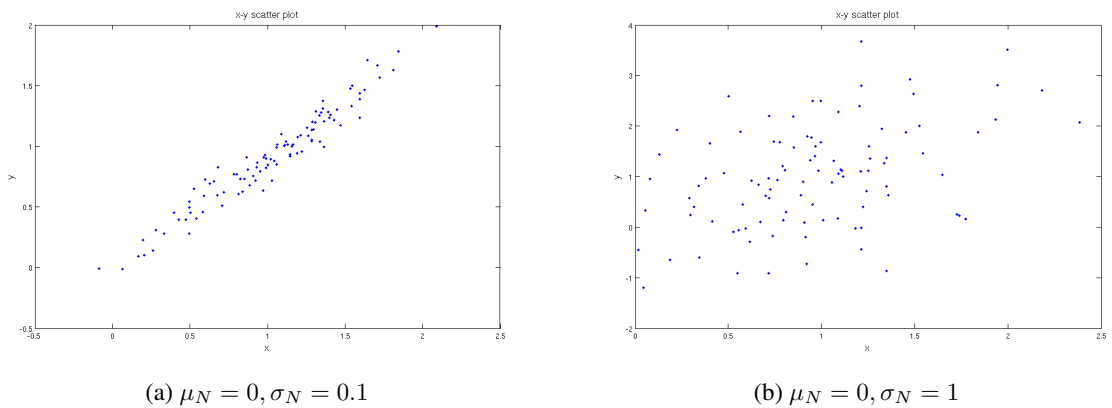


Figure 4.7: 模擬方式三的 XY 散佈圖 II

4.4 數據模擬結果與真實數據之比較

在 4.1 ~ 4.3 節中介紹了三種可能造成振幅衰減的統計模型，並利用數值模擬檢查模擬後的隨機變數相關性。在本節中，將比對真實數據與統計模型的模擬結果，選擇最符合事實的模型做為振幅衰減的統計模型。

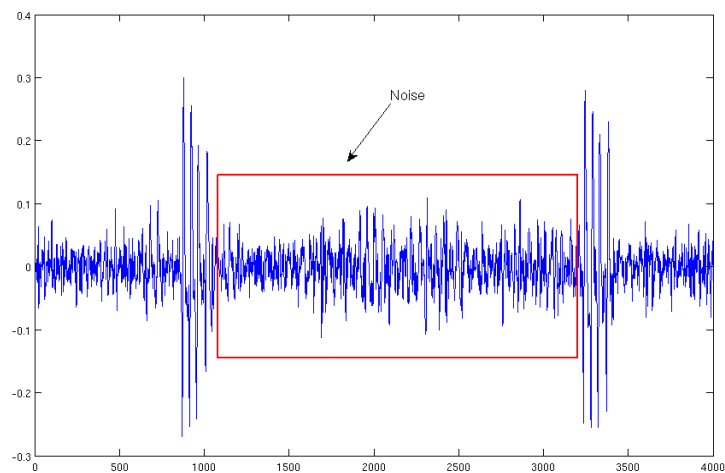


Figure 4.8: 雜訊選取

在模擬方式三會需要加上雜訊，我們自原始訊號中取一段沒有電作電位的訊號代表雜訊，如 Figure 4.8。

雜訊需要經過和訊號相同的前置處理程序，計算該段訊號的平均值與標準差做為模擬方式三的雜訊模擬設定值。三種統計模型中的 X 是代表 units 的第一個振幅，

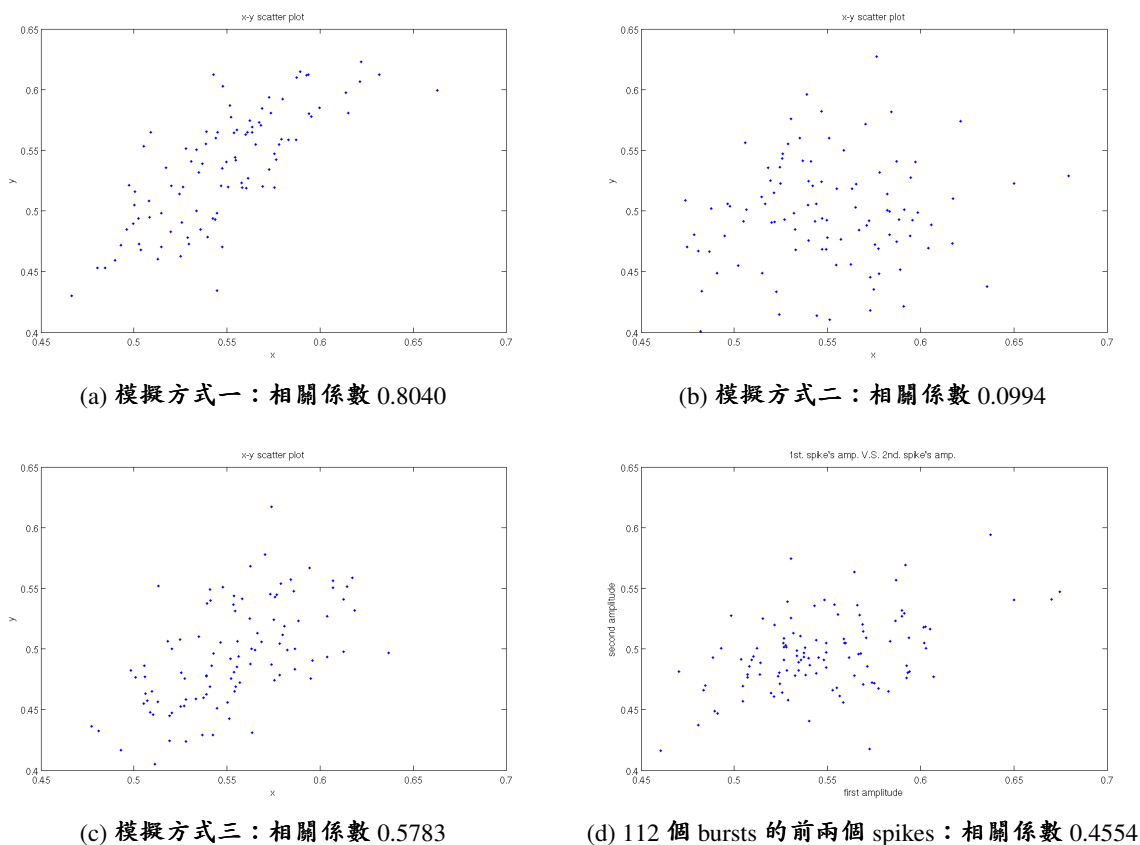


Figure 4.9: 真實數據中的統計模型相關數值表

為了與真實數據比較，在模擬 X 時的設定值即以真實數據中的第一個 units 振幅之平均值與標準差為準。同理，在模擬方式一中的 R 以真實數據中的第一個 units 與第二個 units 之振幅比值做為設定值。在 Table 4.1 中整理出三種模擬方式所需要的數值。Figure 4.9 是三種模擬方式與 112 個 bursts 的散佈圖比較。

	X (第一個 units 振幅)	R ($\frac{\text{第二個 units 振幅}}{\text{第一個 units 振幅}}$)	N (雜訊)
平均值	0.5484	0.9076	-0.0003
標準差	0.0392	0.0627	0.0323

Table 4.1: 前二個 units 振幅平均值與標準差

Figure 4.9 的 (a)、(b)、(c) 三圖，分別利用模擬方式一、二、三，搭配真實訊號得到的數據 (Table 4.1) 所做出來的結果，(d) 圖是真實訊號的結果。比較三種模擬方式最後得到的相關係數，以及觀察散佈圖的分佈情況，決定 (c) 圖與 (d) 圖最為相近，所以認為模擬方式三是造成振幅衰減的統計模型。

第 5 章 結論與展望

在神經訊號中，不同的頻率、振幅或者形式 (spikes 或 bursts)，可能都代表著不同的意義。每一根微電極所紀錄到的細胞數量可能不止一個，而每一個神經細胞所產生的訊號也可能不止一種。如何有效且即時地自動分類訊號，並探測神經細胞個數，在神經訊號研究上是個重要的前置處理。為了達到「即時且自動」的目的，電腦的輔助是不可或缺的，因此我們需要將人眼判斷的標準加以量化。

5.1 結論與探討

當拿到未經處理的原始訊號時，本文依下列步驟處理：

1. 降低取樣頻率。降維的目的是在不影響特徵的情況下減少資料量，以縮短處理的速度。可以使用附錄 A.1 的降維函式，輸入原始的取樣頻率以及最後希望得到的頻率是多少。
2. 濾波。本文所使用的濾波方式是 shrink，是基於白雜訊的假設下來去除雜訊。如果雜訊的分佈不是白雜訊，則應該改用其它的濾波方式。附錄 A.2 是一個簡單的 shrink 函式，詳細使用方式參照附錄。
3. 偵測動作電位。本文使用的偵測方式是門檻探測法 [7] (附錄 A.3)。找出動作電位後再進一步區分為 spikes 與 bursts (附錄 A.4)。
4. 檢驗 bursts。必須檢查在前一個步驟篩選出來的 bursts 是否為同一個神經細胞所產生 (附錄 A.5 與 A.6)。

附錄 A.8 是一個簡單的處理流程，裡面運用了上述所提到的步驟與函式，這些函式是自行定義的，並非 Matlab 的內建函式。

本文探討的 bursts 是一種特別的訊號模式，探針在大腦皮層中的紀錄幾乎只有看到 spikes，沒有看到 bursts；而視丘除了 spikes 之外，bursts 的數量也相當多，出現頻次 (每單位時間出現的次數) 遠高於大腦皮層。我們需要給 bursts 一個定性定量的描述，以便能夠自動化地偵測和區分 spikes 與 bursts。從已知為 bursts 的訊號之中，本文最後歸納出三個可以用來篩選 bursts 的條件。第一篩選條件是 unit 與 unit 之間的時

間隔需介於 1 ~ 4 毫秒內，從第二章我們知道，神經細胞有所謂的絕對不反應期 1 毫秒，以及相對不反應期 4 毫秒，在相對不反應期內，除非細胞受到更大的刺激，否則不會主動再發生反應。而 bursts 是在未受到額外刺激之下，於相對不反應期間主動發生反應，這一點與過去熟知的 spikes 訊號是不同的。第二篩選條件是每一次的振幅減少為前一次的 90%，直到最後一次減少為第一次的 80% 為止。在實驗過程中我們看到一個現象：在已知的 112 個 bursts 訊號中，由三個 units 所組成的 bursts 數量最多，次多的是四個 units，最後是五個 units。

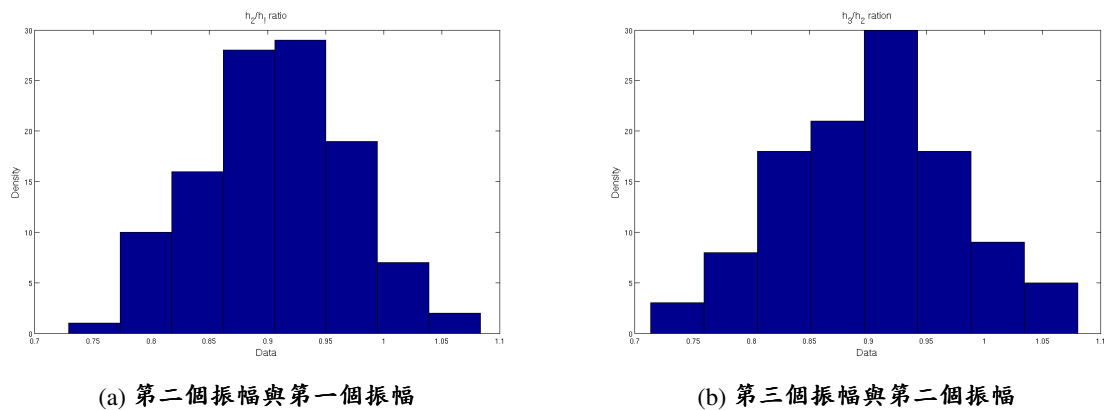


Figure 5.1: bursts 振幅間比直直方圖

從第二篩選條件、Table 3.1 ~ 3.3 以及 Figure 5.1 可以看到，units 間的振幅比值約為平均值 0.9 的常態分佈。假設第一個振幅是 1，則 $0.9^2 = 0.81$ 再考慮雜訊的影響，第三個振幅就相當接近 0.8，也到達停止 bursts 訊號的條件；若是 0.93 的話，則需要到達第四個振幅才能接近停止條件。由於常態分佈的關係，接近 0.9 比率的個數最多，距離 0.9 愈遠則個數愈少，此分佈也說明了為什麼觀察到以三個 units 組成的 bursts 訊號居多。

5.2 檢討與建議

在訊號處理中，去除雜訊是必要的。但是在濾波的過程中，如果濾掉了太多的資訊，將使得 bursts 振幅遞減的特性消失。Figure 5.2 是一個 20KHz 取樣頻率下振幅明顯遞減的 burst 訊號，其在不同程度的濾波後之比較圖。Figure 5.2(a) 是原始的訊號，可以看到有一些細節資訊是我們不需要的，例如 burst 裡的第三個 spike 在高峰處有許多更小的「波」。這些不影響訊號特徵的細節資訊對於訊號而言是雜訊，因

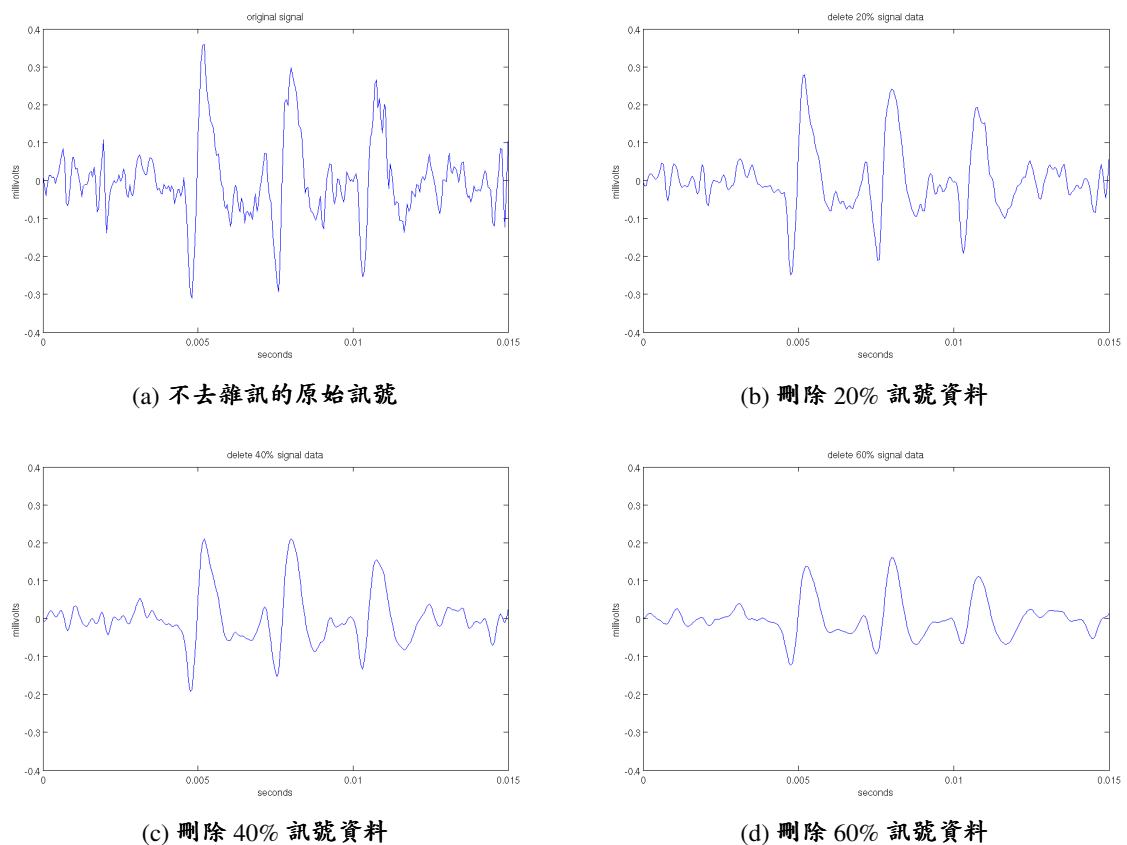
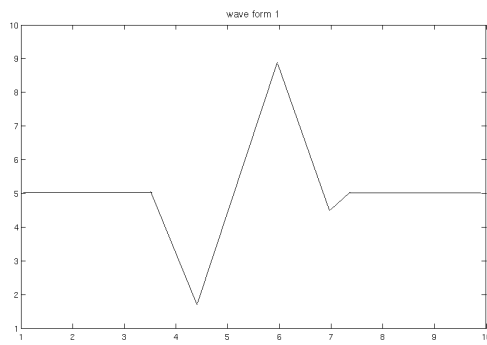


Figure 5.2: 去雜訊比例與振幅比較圖

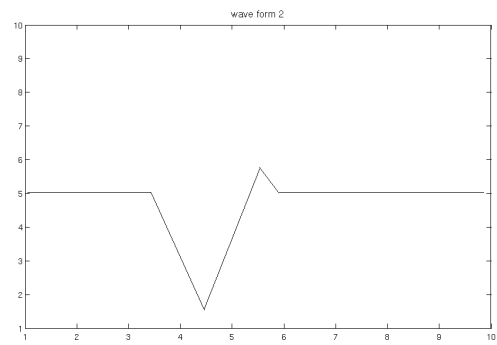
此爲了簡化判斷以及減少錯誤，需要去除雜訊。

在本研究的實驗過程中有注意到一個現象，當濾波刪除掉的資料超過 20% 時，振幅遞減的特徵將不再明顯。從 Figure 5.2 (c) 和 (d) 圖來看，該 burst 訊號經過 40% 或 60% 的濾波後，振幅遞減的特徵已不如 Figure 5.2(a) 和 (b) 圖清楚。由於目前對神經訊號的雜訊組成並不完全明瞭，在濾波時是以白雜訊做爲假設，並不完全符合實際雜訊的分佈情況。未來如果對於神經訊號的雜訊有更多的瞭解，那麼將來在濾波時便可以針對雜訊做適當的處理，如此不但可以去除真正雜訊的影響，並且能夠保留振幅遞減的特徵。目標訊號的特徵愈明顯，對於後續篩選的動作將愈有利。

動作電位的波形可能有很多種，目前知道兩種波形 [4]。Figure 5.3 是兩種波形的示意圖。由於本文所統計的 112 個 bursts 訊號都是波形一的類型，爲了方便分析數據，以及檢驗篩選條件正確與否，程式設計上也是針對波形一來實作，這是需要改進的地方。附錄的程式是以 Matlab 實作，爲了更快看到實驗的結果，我們並沒有很精緻的設計程式。爲了方便起見，有時候會使用平鋪直序的程式設計，以求快速看



(a) 波形一



(b) 波形二

Figure 5.3: 動作電位波形示意圖

到實驗結果。但是以長遠的角度來看，這種方式並不是個好的程式設計方法，最大的缺點之一是程式不易維護，且作者以外的程式設計師不容易看懂程式的內容。未來在判斷以及截取訊號的部份，應該要配合 Matlab 的特性，以及研究訊號的波形與特性使用良好的演算法來增進效率，以及簡化程式。

參 考 文 獻

- [1] S. Murray Sherman. A wake-up call from the thalamus. *nature neuroscience*, volume 4:344–346, 2001.
- [2] 吳蕙稜. 神經元電位訊號雜訊之組成分析. Master's thesis, 國立中央大學, 2008.
- [3] 單維彰. 凌波初步 *First Concepts of Wavelets*. 全華科技, 1999.
- [4] 李易霖. 神經細胞訊號的一種分類方法. Master's thesis, 國立中央大學, 2004.
- [5] 柯正玲. 以點過程分析動作電位與為反應之關聯性. Master's thesis, 國立中央大學, 2009.
- [6] 王錫崗等譯. 人體生理學. 新文京開發出版股份有限公司, 2004.
- [7] 謝博文. 自動分類軟體在動作電位上的研究. Master's thesis, 國立中央大學, 2006.

附錄

附錄是在本文中**使用過**的相關函式，以及一個處理一筆原始訊號的簡單流程。

A.1 降維函式

```
% This function is going to down sampling and normalization signal.
%
% 1. Down sample to 20KHz.
% 2. Normalize signal.
%         newSignal = (oldSignal-mean(oldSignal))/std(oldSignal)
%
% isig: Input signal
% ori_freq: Original signal frequency
% out_freq: Output signal frequency
% option:
%         1: Down sampling only.
%         0: Down sampling and normalization.

function [osig, result] =
    down_sample(isig, ori_freq, out_freq, option);

osig = [];
% 目標維度最好是原始維度的倍數，否則在此函式中暫時不處理。
if mod(ori_freq, out_freq)~=0,
    result = 1;
else
    if option == 0,
        % 計算取樣間隔
        step = ori_freq/out_freq;
        % 自原始訊號中，按照上一步所計算的取樣間隔，重新取樣。
        osig = isig(1:step:end);
        % 訊號標準化。將資料調整成平均值為 0，標準差為 1。
        osig = (osig-mean(osig))/std(osig);
    elseif option == 1,
        step = ori_freq/out_freq;
        osig = isig(1:step:end);
    end
    % 函式無誤回傳值
    result = 0;
end
```

A.2 去雜訊函式

```
% 去雜訊函式
% 這裡使用 shrink 方式濾波，適用於雜訊分佈
% 大致呈高斯分佈的雜訊。
% Shrink 濾波詳情可參照 Figure 4.5
%
% t: 傳入訊號的長度，以秒為單位
% isig: 傳入欲濾波的訊號
% freq: 傳入訊號的頻率

function [osig] = shrink(t, isig, freq, bound);

    time = t;
    % 資料長度
    N = length(isig);
    % 利用 Matlab 內建的傅立葉函式做轉換
    sig = fft(isig,N)/freq;
    % 原始資料為實數，轉換後的數值有虛數，虛數成對
    % 因此只需一半的資料即可。
    spec = sig(1:N/2);
    % 設定欲過濾的閾值
    epslon = bound;
    % 計算欲刪除的資料量
    newSpec = abs(spec);
    newSpec = sort(newSpec);
    delData = sum(newSpec);
    delData = bound*delData;
    % 產生暫存空間
    tmpBuf = 0;
    % 累加排序後的資料，直到到達閾值為止
    for j = 1:1:N/2,
        tmpBuf = tmpBuf + newSpec(j);
        if tmpBuf > delData,
            epslon = newSpec(j);
            break;
        end
    end
end

% 轉換後的數值為複數，可表示為  $re^{\{theta\}}$ 
% 參考 Figure 4.5，shrink 濾波是將每個頻率
% 的強度都降低，以複數平面來看，即是縮短
% 半徑，角度不變。若原始半徑即小於閾值，
% 則以 0 計算之。
for j = 1:1:N,
    if abs(sig(j)) < epslon,
```



```

        sig(j) = 0;
    else
        a = angle(sig(j));
        sig(j) = ( abs(sig(j))-epsilon)*(exp(i*a));
    end
end
end
% 清除記憶體用量大的變數
clear isig;
% 將頻譜傅立葉反轉換回訊號
osig = real(ifft(sig*N/time));

```

A.3 找局部極值函式

本函式使用「中央大學數據分析中心」公開的「HHT MATLAB program」其中一個函式，再依需要做修改。

```

% extrema 函式主要目的為找出超過門檻值的局部極值。
% in_data: 輸入的訊號
% threshold: 門檻值
% flag: 輸出局部極值的
%         位置、數值、是否超過閾值、是否為極大值、是否為極小值
%
function [flag]= extrema(in_data, threshold)
% 資料長度
dsize=length(in_data);
% 為避免記憶體存取費時，先產生一個足
% 夠大的空間
flag = ones(dsize,5);

jj=2;
kk=1;
kk2=1;
while jj < dsize,
    tag = 0;
    % 找局部極大值
    if ( in_data(jj-1)<=in_data(jj) & in_data(jj)>=in_data(jj+1) )
        & in_data(jj) > threshold,
        % 超過閾值標記 1
        if in_data(jj) > threshold,
            flag(kk,3) = 1;
        % 未超過閾值標記 0
        else

```

```

        flag(kk,3) = 0;
    end
    % 記錄相關資訊
    flag(kk,1) = jj;
    flag(kk,2) = in_data (jj);
    flag(kk,4) = 1;
    flag(kk,5) = 0;
    kk = kk+1;
    tag = 1;
end
% 找局部極小值
if ( in_data(jj-1)>=in_data(jj) & in_data(jj)<=in_data(jj+1)
    & tag ~= 1) & in_data(jj) < -threshold,
    % 超過閾值標記 1
    if in_data(jj) < -threshold,
        flag(kk,3) = 1;
    % 未超過閾值標記 0
    else
        flag(kk,3) = 0;
    end
    % 記錄相關資訊
    flag(kk,1) = jj;
    flag(kk,2) = in_data (jj);
    flag(kk,4) = 0;
    flag(kk,5) = 1;
    kk = kk+1;
end
jj=jj+1;
end
% 刪除未使用的空間
flag(kk:end,:) = [];

```

A.4 偵測 spikes 與 bursts 函式

```

% detect 函式用於初步篩選 bursts 與 spikes
% 使用第一條件以及第二條件 (見 4.2 節)
% 做為篩選 bursts 的第一步，其它超過閾
% 值且不符合第一條件的動作電位則為 spikes。
%
% sig: 原始訊號
% extrem_value: 局部極值的相關資訊
% act_freq: 訊號的取樣頻率
%
function [burst, spike] = detect(sig, extrem_value, act_freq)

```

```

% 變數宣告以及初始化
no = 1;
j = 1;
last_index = 0;
L = length(extrem_value);
dt = 1/act_freq;
% 紀錄編號、起點、終點、總時間、spikes 個數
burst = ones(L,5);
amp = zeros(L,10);
spike = ones(L,52);
end_index = 1;
tag_burst = 0;
tag_spike = 0;
count_burst = 0;
count_spike = 0;
ok = 0;
str_pt = 0;
max_count = 0;

% 第一條件: spikes 的時間間隔需介於 1ms~4ms 之間
while j < L,
    % extrem_value 格式: 起點 終點 有無超過門檻 最大 最小
    % 當 min 超過 threshold 時, 視為一個反應開始發生
    if extrem_value(j,5) == 1,
        tag_burst = 0;
        tag_spike = 0;
        ok = 0;
        str_pt = extrem_value(j,1);
        str_amp = 0;
        if j+20 > L,
            break;
        end
        for k = (j+1):1:(j+20),
            if extrem_value(k,4) == 1,
                if extrem_value(k,1)-str_pt < 32,
                    str_amp = extrem_value(k,2)-extrem_value(j,2);
                    ok = 1;
                    break;
                end
            end
        end
    end
    if ok == 1,
        tmp = j+20;
        if tmp > L,
            break;
        else
            jplus = 0;
            % 4ms = 80 points, burst 至少要在

```

```

% 4ms 內有第二個 spike 發生
for k = (j+1):1:(tmp-1),
    if extrem_value(k,5)==1
        & extrem_value(k+1,4)==1,
            if extrem_value(k,1)-str_pt < 80
                & extrem_value(k,1)-str_pt > 20,
                    tag_burst = 1;
            end
        end
    end
    if tag_burst == 1,
        break;
    end
end
% 若 4ms 內都沒有找到下一個 spike，則此次視為 spike
if tag_burst == 0,
    tag_spike = 1;
end
% burst
if tag_burst == 1,
    tmp = j+20;
    count = 0;
    stop_sig = 0;
    pre = extrem_value(j,1);
    for k = j:1:tmp,
        if extrem_value(k,5) == 1,
            if k == j,
                for kk = (k+1):1:tmp,
                    if extrem_value(kk,4)==1
                        & extrem_value(kk,1)-extrem_value(k,1)<40,
                            count = 1;
                            amp(count_burst+1,count) =
                                extrem_value(kk,2)-extrem_value(k,2);
                            jplus = kk;
                            break;
                        end
                    end
                end
            elseif stop_sig==1,
                break;
            elseif extrem_value(k,1)-pre<80
                & extrem_value(k,1)-pre>20,
                    for kk = (k+1):1:tmp,
                        if extrem_value(kk,4)==1
                            & extrem_value(kk,1)-extrem_value(k,1)<40,
                                count = count+1;
                                amp(count_burst+1,count) =
                                    extrem_value(kk,2)-extrem_value(k,2);
                                if amp(count_burst+1,count)
                                    <= str_amp,

```

```

        jplus = kk;
    else
        count = count-1;
        stop_sig = 1;
    end
    break;
end
end
pre = extrem_value(k,1);
else
    break;
end
else
    continue;
end
end
if count > 1,
    sig_str = extrem_value(j,1)-9;
    sig_end = extrem_value(kk,1)+22;
    % burst 發生時間長度要合理
    if sig_end-sig_str>20*(count-1)
        & sig_end-sig_str<80*count,
            count_burst = count_burst+1;
            burst(count_burst,1) = count_burst;
            burst(count_burst,2) = sig_str;
            burst(count_burst,3) = sig_end;
            burst(count_burst,5) = count;%
            if max_count < count,
                max_count = count;
            end
            j = j+1;
        else
            amp(count_burst+1,:) = [];
            j = j+1;
        end
    else
        j = j+1;
    end
end
% 如果是 spike
elseif tag_spike == 1,
    count_spike = count_spike+1;
    sig_str = extrem_value(j,1)-19;
    sig_end = extrem_value(j,1)+32;
    spike(count_spike,:) = sig(sig_str:sig_end);
    j = j+2;
% 如果發生了不是 burst 也不是 spike 的情況
% 代表程式有問題。
else

```

```

        fprintf('Detecting is wrong.\n');
        return;
    end
end
else
    j = j+1;
end
else
    j = j+1;
end
end
% 刪除未使用到的空間
burst((count_burst+1:end), :) = [];
spike(count_spike+1:end, :) = [];
amp((count_burst+1:end), :) = [];
amp(:, max_count+1:end) = [];

% 第二條件：振幅要漸減
burst2 = zeros(length(burst)); % 編號 起點 終點 個數
burst3 = [];
amp3 = [];
fprintf('sensor1:%03d\n', length(burst));
for j = 1:1:count_burst,
if burst(j,5) < 3,
    burst2(j) = 1;
else
    for jj = 1:1:max_count-1,
        if amp(j, jj) ~= 0,
            % 還沒有趨於平緩，設定在 0.9+-0.06
            if jj+2<=max_count & amp(j, jj+1) ~= 0 & amp(j, jj+2) ~= 0,
                if amp(j, jj+1)/amp(j, jj) < 0.78
                    | amp(j, jj+1)/amp(j, jj) > 1.02,
                        burst2(j) = 1;
                        break;
                    end
                end
            % 趨於平緩，設定在 0.9+-0.12
            elseif jj+2<=max_count & amp(j, jj+1) ~= 0 & amp(j, jj+2) == 0,
                if amp(j, jj+1)/amp(j, jj) < 0.78
                    | amp(j, jj+1)/amp(j, jj) > 1.02,
                        burst2(j) = 1;
                        break;
                    end
                end
            % 後面是0，代表自己是最後一個 spike
        else
            % 按推論，最後一個應該與第一個相差 0.8 左右
            if amp(j, jj)/amp(j, 1) < 0.74 | amp(j, jj)/amp(j, 1) > 0.86,
                burst2(j) = 1;
                break;
            end
        end
    end
end
end

```

```

                end
            end
        else
            break;
        end
    end
end
end
if burst2(j) == 0,
    burst3 = [burst3; burst(j,:)];
    amp3 = [amp3; amp(j,:)];
end
end
a = size(burst3);
burst = burst3;

```

A.5 檢查 bursts 的相對誤差函式

```

% err_calc 函式即為第三條件的判斷
% 將已符合第一條件以及第二條件的訊號
% 傳入此函式中篩選，最後回傳 bursts 訊號
%
% b_: 符合第一以及第二條件的訊號
% sig: 原始訊號
% thres: 門檻值
%
function [wave_form,unitCount] = err_calc(b_, sig, thres);

% 計算誤差?
j = 1;
jj = 1;
L = size(b_);
L = L(1);
wave_form = zeros(L,26);
unitCount = zeros(L,1);

while j <= L,
    tmp_no = b_(j,5);
    tmp_str = b_(j,2);
    tmp_end = b_(j,3);
    % 將訊號一個一個取出，放在暫存空間，方便使用
    tmp = sig(tmp_str:tmp_end+10);
    % 訊號長度需合理
    if length(tmp) < 500,
        % 訊號傳入 seperate_burst 函式做分割與計算

```

```

% seperate_burst 函式回傳該訊號的相對誤差、
% 代表波形、spikes 振幅、spikes 個數
[err form emp count] = seperate_burst(tmp, j, thres, tmp_no);
% 相對誤差不可大於 50%，此處設定 50%，是參考
% 已確定為 bursts 訊號的計算結果。
if sum(err)/(tmp_no-1)>0 & sum(err)/(tmp_no-1)<0.5,
    wave_form(jj,:) = form;
    unitCount(jj) = count;
    jj = jj+1;
% 若相對誤差大於 50%，則輸出提示訊息
% 且忽略該訊號。
else
    fprintf(['burst_no. ', num2str(j)
            , ' deleted because of error '
            , num2str(sum(err)/(tmp_no-1)), '.\n']);
end
end
    j = j+1;
end
% 刪除不使用的空間
wave_form(jj:end,:) = [];
% wave form 的 prototype 做正規化以及對齊
wave_emp = max(wave_form(1,:))-min(wave_form(1,:));
figure, hold on,
wave_form2 = zeros(L,24);
matSize = size(wave_form);
% 輸出提示訊息
fprintf('total bursts: %03d\n',a(1));
for j=1:1:matSize(1),
    if (wave_form(j,2)-wave_form(j,1)) == 1;
    else
        tmpBuf = min(wave_form(j,:));
        k = 1;
        for k = 1:1:26,
            if wave_form(j,k) == tmpBuf,
                tmpBuf2 = wave_form(j,k-3:k+20);
                tmpBuf2 = tmpBuf2-(tmpBuf-min(wave_form(1,:)));
                plot(tmpBuf2);
                wave_form2(j,:) = tmpBuf2;
                break;
            end
        end
    end
end
end
hold off;

```


A.6 切割 bursts 為 spikes 回傳相對誤差以及代表波形函式

```
% seperate_burst 函式將適當切割 bursts 訊號
% 回傳其相對誤差、代表波形...等相關資訊。
%
% sig: burst 訊號
% no: burst 的編號
% thres: 門檻值
% tol_no: spikes 個數
%
function [err,prototype,b_emp,b_no]
    = seperate_burst(sig,no,thres,tol_no);

    % 變數宣告與初始化
    wave_form_length = 26;
    s = length(sig)-1;
    b_no = 0;
    b_emp = [];
    b_s = [];
    tmp_str = 0;
    tmp_end = 0;
    tmp = zeros(1,26);
    x = 1;
    err3 = -1*ones(1,3);
    b_base = [];
    prototype = [];
    last_local_min = 0;
    for j = 2:1:s,
        if sig(j)<=sig(j-1) & sig(j)<=sig(j+1) & sig(j)<-thres,
            j_tmp = j+26;
            if j_tmp > length(sig),
                j_tmp = length(sig);
            end
            if max(sig(j:j_tmp))>thres,
                x = 1;
                if b_no == 0,
                    last_local_min = j;
                else
                    if (j-last_local_min) < 20,
                        j = j+1;
                        continue;
                    else
                        last_local_min = j;
                    end
                end
            end
            b_no = b_no+1;
            tmp_str = j-4;
```

```

tmp_end = tmp_str+25;
if tmp_end > length(sig),
    tmp_end=length(sig);
end
xdd = max(sig(tmp_str:tmp_end));
% 如果是個正常的動作電位，max 應大於 threshold，
% 否則視為一個異常的 noise
if xdd > thres,
    tmp(x) = sig(tmp_str);
    x = x+1;
    % 檢查資料長度，使 spikes 長度統一
    if tmp_end > s,
        tmp_end = s;
        tmp_s = tmp_end-tmp_str;
        b_s(:, ((2+tmp_s):end)) = [];
    end
    % 取出 spike 以及計算 amplitude
    for k = (tmp_str+1):1:(tmp_end-1),
        if sig(k)>=sig(k-1) & sig(k)>=sig(k+1)
            & sig(k)>thres,
                b_emp = [b_emp; (sig(k)-sig(j))];
            end
        tmp(x) = sig(k);
        x = x+1;
    end
    tmp(x) = sig(tmp_end);
    aaa = (max(sig(tmp_str:tmp_end))
        -min(sig(tmp_str:tmp_end)));
    aaa = b_emp(1)/aaa;
    % 左右各移一點，找誤差最小的位置。
    if b_no > 1,
        err3(1) = sqrt(sum((b_base-aaa*
            sig(tmp_str-1:tmp_end-1)') .^2))
            *100/sqrt(sum(b_base.^2));
        err3(2) = sqrt(sum((b_base-aaa*
            sig(tmp_str:tmp_end)') .^2))
            *100/sqrt(sum(b_base.^2));
        err3(3) = sqrt(sum((b_base-aaa*
            sig(tmp_str+1:tmp_end+1)') .^2))
            *100/sqrt(sum(b_base.^2));
        xd = min(err3);
        % err3
        for x = 1:1:3,
            if err3(x) == xd,
                tmp = sig(tmp_str+x-2:tmp_end+x-2)';
                break;
            end
        end
    end
end

```

```

elseif b_no == 1,
    b_base = tmp;
end
a = size(b_emp);
if a(1) == b_no,
    b_s = [b_s; tmp];
    j = tmp_end+1;
else
    j = j+1;
    b_no = b_no-1;
end
j = tmp_end;
% 如果 max 太小，代表只是個比較大的 noise，則不再做
end
end
end
if b_no == tol_no,
    break;
end
end

matSize = size(b_s);
n = matSize(1);
if n == tol_no,
    t = (1:1:matSize(2));
    b_base = b_s(1,:);
    base_emp = b_emp(1);
    err = [];
    % 誤差估計：以 2-norm 做為測量方式計算相對誤差
    % 儲存 bursts 的第一個 spike 波形做為該 bursts 的代表波形
    if n == 3,
        b_s(2,:) = (base_emp/b_emp(2))*b_s(2,:);
        b_s(2,:) = b_s(2, :)-(b_s(2,5)-b_base(5));
        b_s(3,:) = (base_emp/b_emp(3))*b_s(3,:);
        b_s(3,:) = b_s(3, :)-(b_s(3,5)-b_base(5));
        err = [err sum(abs((b_base-b_s(1,:))))/sum(abs(b_base))];
        err = [err sum(abs((b_base-b_s(2,:))))/sum(abs(b_base))];
        err = [err sum(abs((b_base-b_s(3,:))))/sum(abs(b_base))];
        prototype = b_base;
    elseif n == 2,
        b_s(2,:) = (base_emp/b_emp(2))*b_s(2,:);
        b_s(2,:) = b_s(2, :)-(b_s(2,5)-b_base(5));
        err = [err sum(abs((b_base-b_s(1,:))))/sum(abs(b_base))];
        err = [err sum(abs((b_base-b_s(2,:))))/sum(abs(b_base))];
        prototype = b_base;
    elseif n == 4,
        b_s(2,:) = (base_emp/b_emp(2))*b_s(2,:);
        b_s(2,:) = b_s(2, :)-(b_s(2,5)-b_base(5));

```

```

        b_s(3,:) = (base_emp/b_emp(3))*b_s(3,:);
        b_s(3,:) = b_s(3, :)-(b_s(3,5)-b_base(5));
        b_s(4,:) = (base_emp/b_emp(4))*b_s(4,:);
        b_s(4,:) = b_s(4, :)-(b_s(4,5)-b_base(5));
        err = [err sum(abs((b_base-b_s(1, :))))/sum(abs(b_base))];
        err = [err sum(abs((b_base-b_s(2, :))))/sum(abs(b_base))];
        err = [err sum(abs((b_base-b_s(3, :))))/sum(abs(b_base))];
        err = [err sum(abs((b_base-b_s(4, :))))/sum(abs(b_base))];
        prototype = b_base;
elseif n == 5,
    b_s(2,:) = (base_emp/b_emp(2))*b_s(2,:);
    b_s(2,:) = b_s(2, :)-(b_s(2,5)-b_base(5));
    b_s(3,:) = (base_emp/b_emp(3))*b_s(3,:);
    b_s(3,:) = b_s(3, :)-(b_s(3,5)-b_base(5));
    b_s(4,:) = (base_emp/b_emp(4))*b_s(4,:);
    b_s(4,:) = b_s(4, :)-(b_s(4,5)-b_base(5));
    b_s(5,:) = (base_emp/b_emp(5))*b_s(5,:);
    b_s(5,:) = b_s(5, :)-(b_s(5,5)-b_base(5));
    err = [err sum(abs((b_base-b_s(1, :))))/sum(abs(b_base))];
    err = [err sum(abs((b_base-b_s(2, :))))/sum(abs(b_base))];
    err = [err sum(abs((b_base-b_s(3, :))))/sum(abs(b_base))];
    err = [err sum(abs((b_base-b_s(4, :))))/sum(abs(b_base))];
    err = [err sum(abs((b_base-b_s(5, :))))/sum(abs(b_base))];
    prototype = b_base;
elseif n == 6,
    b_s(2,:) = (base_emp/b_emp(2))*b_s(2,:);
    b_s(2,:) = b_s(2, :)-(b_s(2,5)-b_base(5));
    b_s(3,:) = (base_emp/b_emp(3))*b_s(3,:);
    b_s(3,:) = b_s(3, :)-(b_s(3,5)-b_base(5));
    b_s(4,:) = (base_emp/b_emp(4))*b_s(4,:);
    b_s(4,:) = b_s(4, :)-(b_s(4,5)-b_base(5));
    b_s(5,:) = (base_emp/b_emp(5))*b_s(5,:);
    b_s(5,:) = b_s(5, :)-(b_s(5,5)-b_base(5));
    b_s(6,:) = (base_emp/b_emp(6))*b_s(6,:);
    b_s(6,:) = b_s(6, :)-(b_s(6,5)-b_base(5));
    err = [err sum(abs((b_base-b_s(1, :))))/sum(abs(b_base))];
    err = [err sum(abs((b_base-b_s(2, :))))/sum(abs(b_base))];
    err = [err sum(abs((b_base-b_s(3, :))))/sum(abs(b_base))];
    err = [err sum(abs((b_base-b_s(4, :))))/sum(abs(b_base))];
    err = [err sum(abs((b_base-b_s(5, :))))/sum(abs(b_base))];
    err = [err sum(abs((b_base-b_s(6, :))))/sum(abs(b_base))];
    prototype = b_base;
end
else
    prototype = (1:1:26);
    err = 0;
    b_emp = 0;
    b_no = 0;

```

```
end
```

A.7 PCA 分析以及 K-means 分群

```
% pcaAndKmeans 函式將對資料矩陣做 PCA 分析，PCA
% 分析後的結果再做 K-means 分群，K-means 分群可指
% 定欲分群的數目以及分群的起始點。
%
% spikeLong: 訊號長度
% V: 資料矩陣
% pc1 = 第一主成份
% pc2 = 第二主成份
% groups = 資料分群個數
% ini = 指定 k-means 初始位置:
% Y = 主成份
% lemnda = 特徵值
function [Y,lemnda,a,b,c,d,s,s2,s3,s4,id] = pcaAndKmeans(spikeLong,V,pc1,pc2,groups,ini)

% 若只有傳入資料矩陣一個變數
% 自動設定第一、第二主成份
% 以及分群個數
if nargin == 1
    pc1 = 1 ;
    pc2 = 2 ;
    groups = 5;
end
[n,m] = size(V) ;
% 後續需要原始矩陣，先存起來
V2 = V;
% 資料矩陣V，每行減去其平均值
for j = 1 : m
    V(:,j) = V(:,j) - mean( V(:,j) ) ;
end
% 計算共變異數矩陣
for i = 1 : m
    for j = 1 : m
        A(i,j) = ( V(:,i)' * V(:,j) ) / n ;
        if i ~= j
            A(j,i) = A(i,j) ;
        end
    end
end
end
% 利用內建函式 eig，將共變異數矩陣 A 對角化
% P:eigenvector, D:eigenvalue
[P,D] = eig(A) ;
```

```

% 儲存最大的特徵值
lemda = flipud( diag(D) ) ;
% 最大特徵值相對應的特徵向量
Y = V*fliplr(P);
%
kmeansData = Y(:,1:2);
if nargin == 6
    [a b c d] = kmeans(kmeansData,groups,'Start',ini);
else
    [a b c d] = kmeans(kmeansData,groups);
end
% 繪製第一主成份以及第二主成份的散佈圖
% 每一群以不同顏色做爲代表，以利區別。
figure, plot( Y(:,pc1) , Y(:,pc2) , '.' ) ;
figure,hold on,
for j=1:1:groups,
    if j==1,
        plot(kmeansData(a==j,1),kmeansData(a==j,2),'r. ');
    elseif j==2,
        plot(kmeansData(a==j,1),kmeansData(a==j,2),'g. ');
    elseif j==3,
        plot(kmeansData(a==j,1),kmeansData(a==j,2),'b. ');
    elseif j==4,
        plot(kmeansData(a==j,1),kmeansData(a==j,2),'c. ');
    elseif j==5,
        plot(kmeansData(a==j,1),kmeansData(a==j,2),'m. ');
    elseif j==6,
        plot(kmeansData(a==j,1),kmeansData(a==j,2),'y. ');
    else
        plot(kmeansData(a==j,1),kmeansData(a==j,2),'k. ');
    end
    hold on,
end
hold off;
% 選中心點做爲該群的代表點
% 中心點對應的波形即爲該
% 群的代表波形
s = zeros(groups,spikeLong);
tmp = 1000*ones(groups,1);
id = zeros(groups,1);
dTmp = d;
for j = 1:1:groups,
    tmp = 1000*ones(groups,1);
    for k = 1:1:length(V),
        if dTmp(k,j)<tmp(j),
            tmp(j) = dTmp(k,j);
            id(j) = k;
        end
    end
end

```

```

end
end
for j = 1:1:groups,
s(j,:) = V2(id(j),:);
end

```

A.8 主程式—處理流程

此處為一個簡單的原始訊號資料處理流程，利用前面幾個小節所提供的函式來處理。

```

% 初始化環境
clc;
clear all;
close all;
% 讀入資料檔。檔案格式為文字檔，只需要 y 軸數值，一行一個數字
% 以此處為例，是讀入檔名為「2010_0116_mice_RTSEL_all.txt」的文字
檔
ori = load('2010_0116_mice_RTSEL_all.txt');
% 原始資料為 40KHz，降至 20KHz
[ori result] = down_sample(ori,40000,20000,1);
% 每個資料點間隔的時間大小
dt = 1/20000;
% 整筆訊號的長度=資料點個數 x 時間間隔
l = length(ori)*dt;
% 確認資料長度是偶數，以便傅利葉轉換
if mod(length(ori),2) ~= 0,
    ori(end) = []; % 刪掉最後一筆資料
    l = length(ori)*dt;
end

% 訊號濾波去雜訊。濾波門檻設 0.2，表示去除 20% 的訊號
shrink_threshold = 0.2;
ori = shrink(l,ori,20000,shrink_threshold);
% 設定截取動作電位的門檻值為 0.7，表示電位差超過 0.7 mV 才抓取訊號
thres_slp_ori = 0.7;
% 標記超過門檻值，且為局部極大或極小值的時間點
[EV_slp_ori] = extrema(ori, thres_slp_ori);
L_slp_ori = length(EV_slp_ori);
% 偵測動作電位，以第一條件和第二條件來區分 spikes 與 bursts
% 將兩種訊號分別存放
[b_slp_ori, s_slp_ori] = detect(ori,EV_slp_ori, 20000);

```

```
% 計算符合「時間間隔」以及「斜率下降」的準 bursts 的相對誤差
% 若誤差符合，則回傳 bursts 代表波形；若EV_slp_ori, 不符合，則去除
[wave_form, unitCount] = err_calc(b_slp_ori, ori, thres_slp_ori);
% 繪製 PCA 散佈圖。不使用 kmeans 分群。
pcaAndKmeans(s_slp_ori,1,2);
```