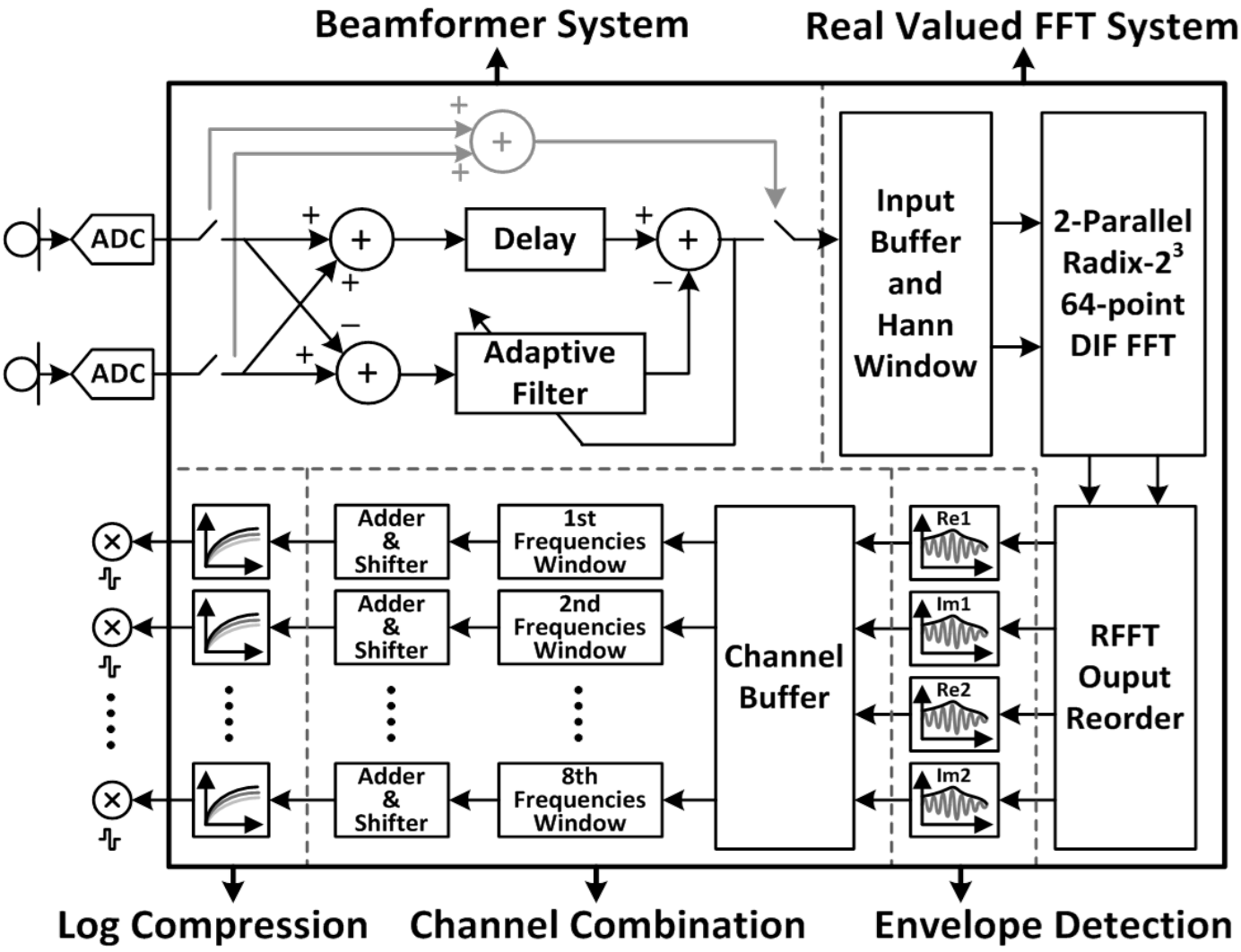
深層類神經網路在語音訊號處理之應用

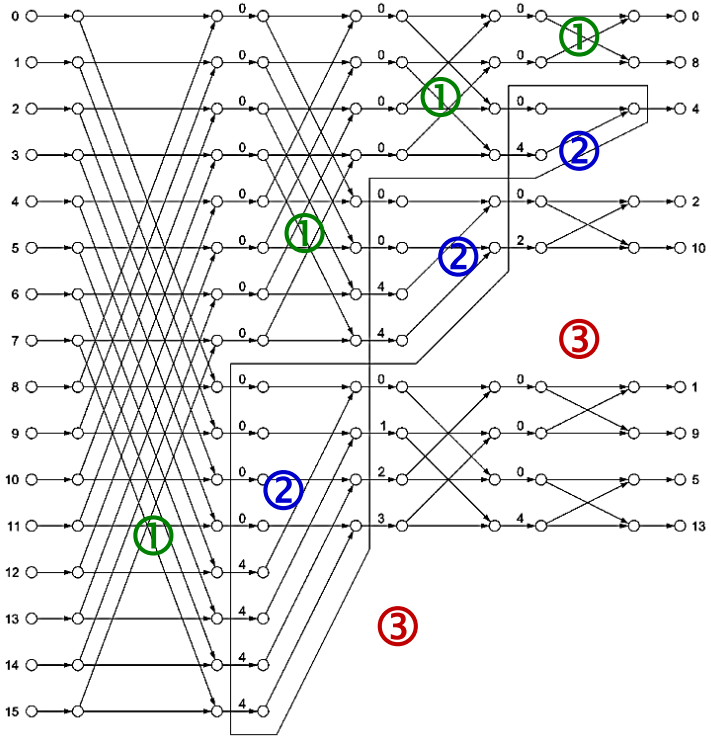
電子工程學系 李諭奇

Proposed Acoustic Digital Signal Processor



Specific Algorithm for Real-Valued FFT

- Three properties to simplify FFT algorithm
  - ① $\mathrm{Im}(\mathrm{x}[n]) = 0$: real-valued data before the first rotation
  - ② Calculation of twiddle factors
  - ③ $\mathrm{X}[\mathrm{N} - \mathrm{k}] = \mathrm{X}^*[\mathrm{k}]$: redundant output



$$\mathbf{Im}(\boldsymbol{x}[\boldsymbol{n}]) = \mathbf{0}$$

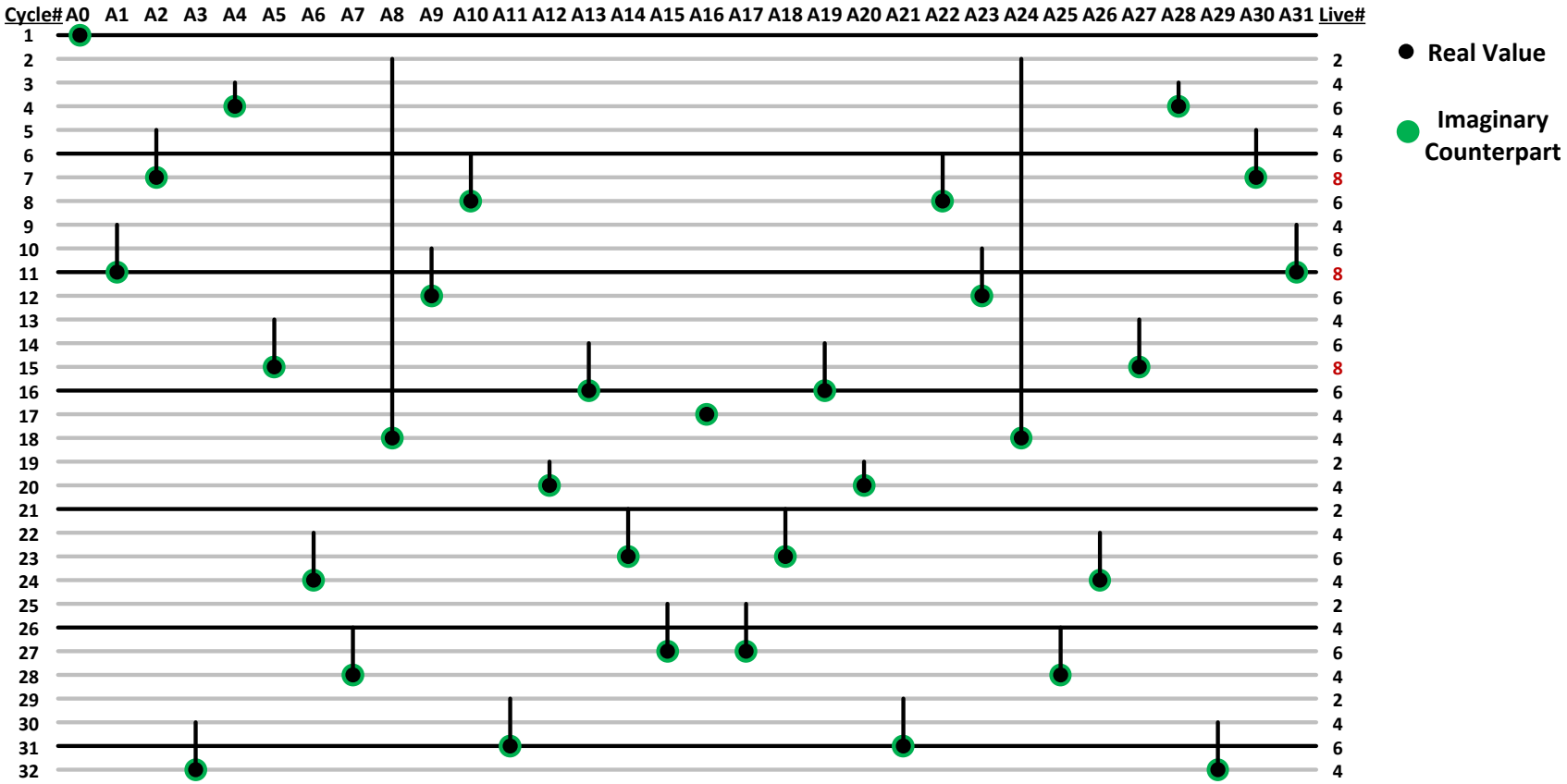
$$\boldsymbol{X_2[i+N/2] = X_1[i+N/2] \cdot e^{-j\frac{2\pi}{N}i}}$$
$$\boldsymbol{+X_1[i+3N/4] \cdot e^{-j\frac{2\pi}{N}(i+N/4)}}$$

$$\boldsymbol{X_2[i+N/2] = \{X_1[i+N/2] - j \cdot X_1[i+3N/4]\} \cdot e^{-j\frac{2\pi}{N}i}}$$

$$\boldsymbol{X[N-k] = X^*[k]}$$

Lifetime Analysis on RFFT Output

- Calculate the minimum number of registers needed to reorder RFFT output
- Linear lifetime chart 64 registers ➡ 8 registers (min. number)



Channel Combination to Eight Channels

- Eight-channel combining method
  - Greenwood function

$$f = \int_0^x \Delta f_{cb} dx = A(10^{\partial x} - k)$$

x: Position of hair cells

f: Corresponding frequency

- FFT bins with corresponding frequencies

Channel No.	FFT bin	Frequencies	Center Freq.
1	1-2	375Hz-750Hz	562.5Hz
2	2-4	750Hz-1500Hz	1125Hz
3	4-6	1,500Hz-2,250Hz	1875Hz
4	6-8	2,250Hz-3,000Hz	2625Hz
5	8-11	3,000Hz-4,125Hz	3562.5Hz
6	11-15	4,125Hz-5,625Hz	4875Hz
7	15-22	5,625Hz-8,250Hz	6937.5Hz
8	22-31	8,250Hz-11,625Hz	9937.5Hz

Future Work

- 未來將建立一個多層次模型來分析不同頻率的語音特徵，在音頻處理方面得到更好的結果，並以深層類神經網路建立自動消除噪音學習機制。

- 與現有發展之人工耳蝸的系統進行整合，燒錄到FPGA板上測試，觀察實際執行結果與後續晶片整合。