

A Twin-8T SRAM Computing-In-Memory Architecture With New Structure Sense Amplifier Application For CNN-Based AI Processor Accelerator

利用一對 8 顆電晶體的靜態隨機存取記憶體進行記憶體內運算和新型感測器應用於卷積神經網路的人工智慧加速器

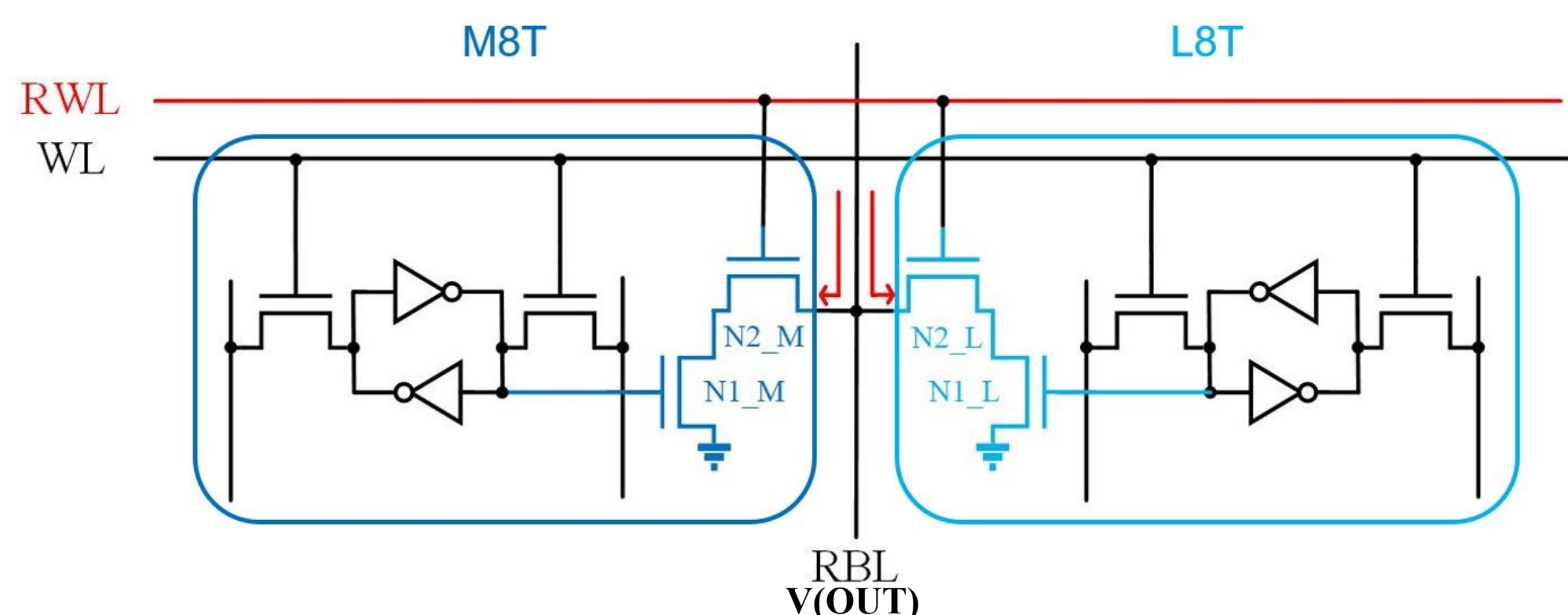
組別 :A29 指導教授 :張孟凡 Mentor :鐘彥麟 組員 :林宸熏、陳禾育

ABSTRACT

在 von Neumann 結構下，電腦、手機等電子儀器在處理各種運算時都需要先將資料從記憶體搬至 CPU 再做運算，碰到需要計算大量資料時，如卷積神經網路 (CNN) 中常使用到的乘積累加運算 (MAC Operation)，便會花費大量的時間和能量在資料傳遞上。為了改善此種情況，在這次的專題中，我們採用 Twin-8T 的靜態隨機存取記憶體 (SRAM) 架構，將 CPU 執行邏輯運算的功能，轉移至記憶體內，並將此記憶體內運算 (Computing In Memory, CIM) 應用於 CNN 的加速，此外也提出一個新的感測放大器 (Sense Amplifier, SA) 架構來增加我們讀取 CIM 結果的成功率。本專題採用製程：CIC018、最高操作電壓：1.8V。

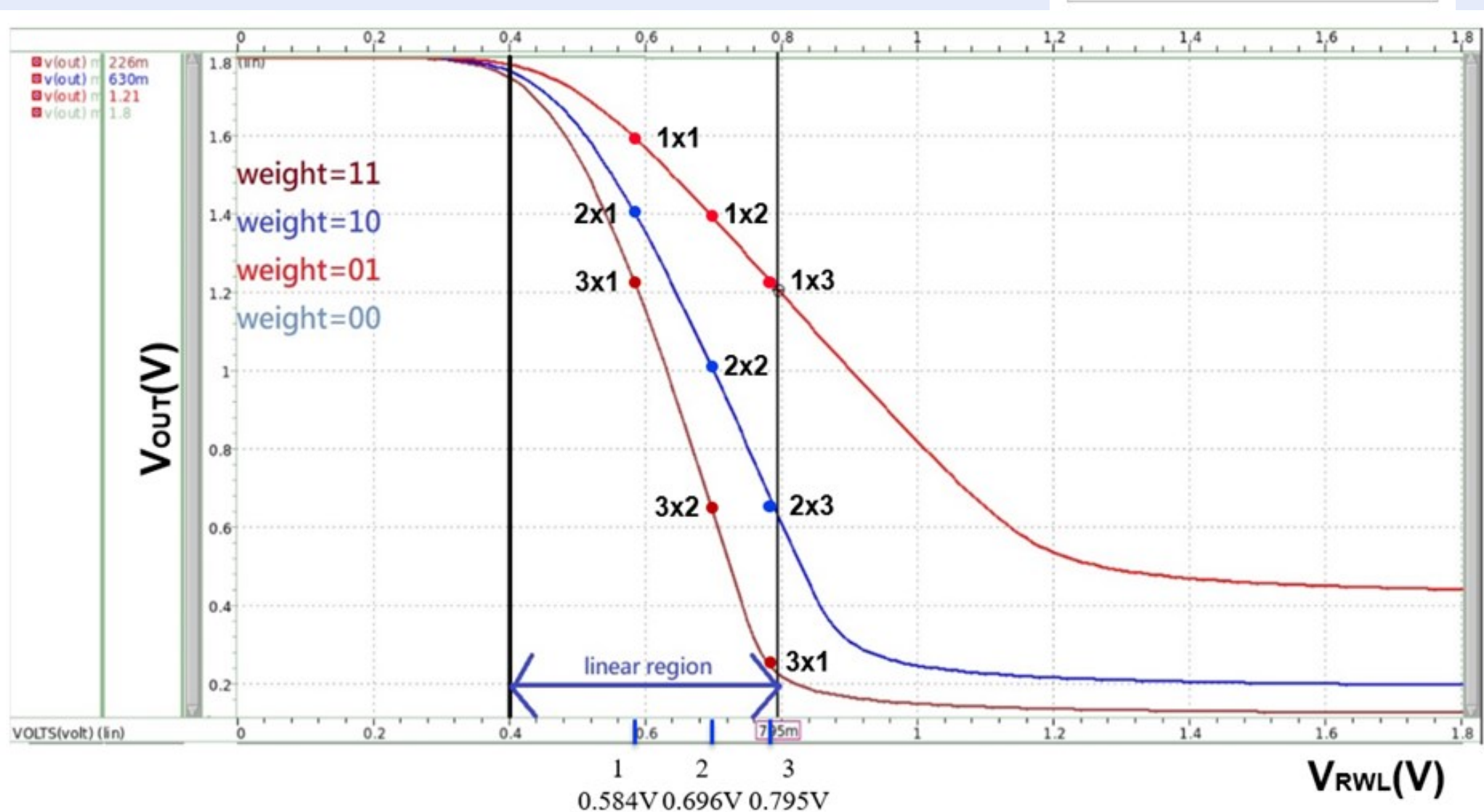
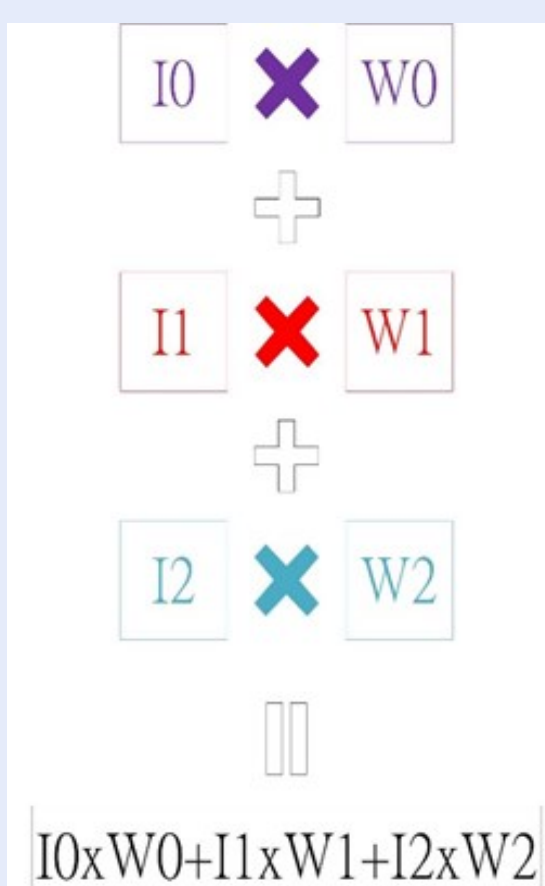
Twin-8T Cell :

TWIN-8T cell 的優點在於一次可以儲存 2-bit 的值，相較一般只能存 1-bit 的 SRAM Cell，能一次做比較大量的 CIM 運算，此外，Cell 左右兩邊都為 8T-SRAM Cell，故具有防止 Read-Disturb 的能力，對於我們在做 CIM 時，能提供更大的穩定性。



MAC Operation :

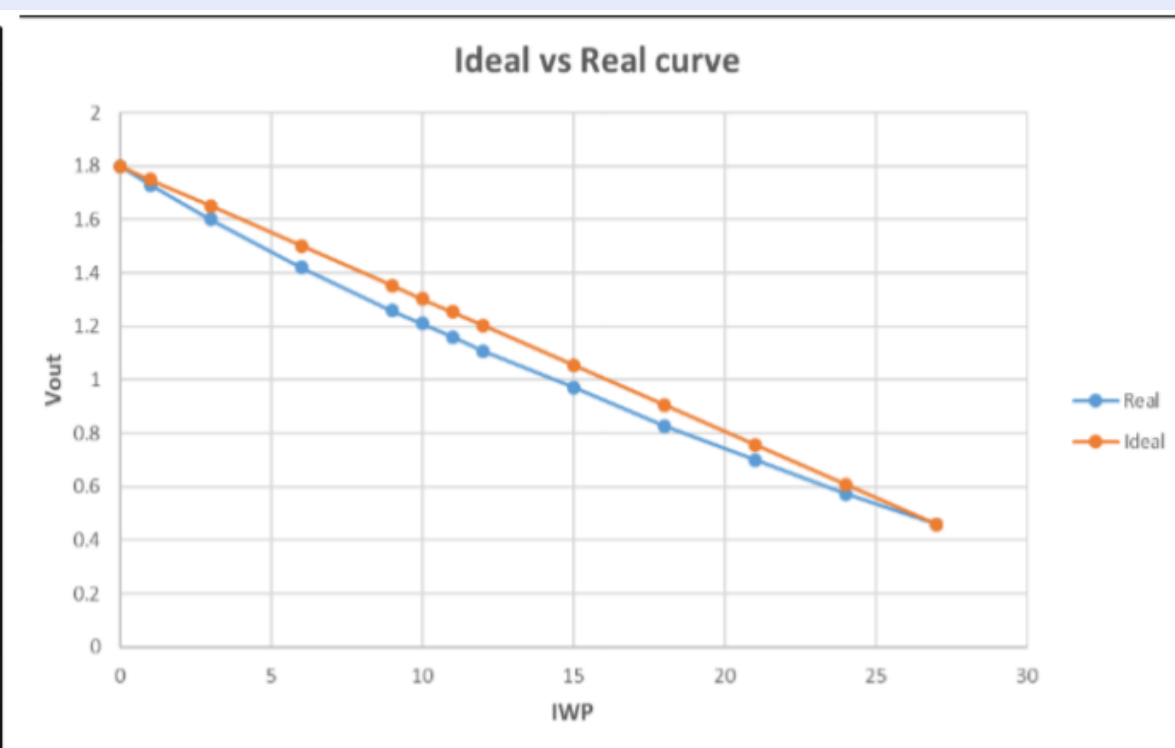
乘加累積的動作是可以一次把許多相乘的結果相加在一起的動作。由右圖來解釋即為若 W 為原本存在記憶體中的值，再來當我有輸入 I 分別灌進 W 做相乘時，能先各自做完乘的動作後，便是一個動作內便把所有相乘出來的值相加起來得到輸出。由於我們做的是 2-bit 的輸入，故我們需要找出分別代表 0、1、2、3 的輸入電壓值，故我們透過 Hspice DC sweep 分析，掃 RWL (輸入電壓) 找出對應到的輸入電壓值。



MAC Operation Result :

此表為不同 MAC 值下對應的模擬與理想 V(OUT)，而繪成圖後可得附圖，可看出理想與模擬上誤差度很低，平均約只有 5.32% 的誤差，整個 MAC operation 有足夠線性的表現。

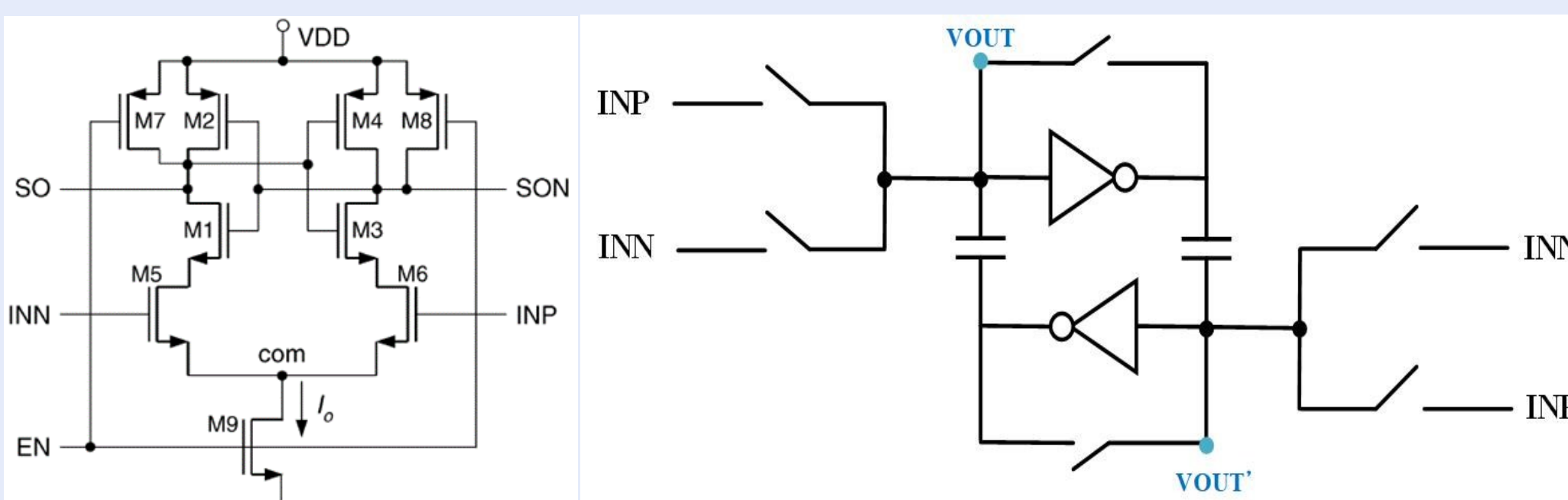
MAC	Vout-R(avg.)(V)	Vout-I(V)	Error(%)
0	1.8	1.80	0
1	1.73	1.75	1.16
3	1.6	1.65	3.10
6	1.42	1.50	5.47
9	1.26	1.35	6.90
10	1.21	1.30	7.19
11	1.16	1.25	7.50
12	1.1075	1.20	8.05
15	0.9725	1.06	7.87
18	0.828	0.91	8.68
21	0.701	0.76	7.49
24	0.5735	0.61	5.81
27	0.46	0.46	0



New Structure Sense Amplifier :

一開始我們嘗試使用傳統 latch type SA (如下左圖) 作為分辨運算結果的感測放大器，但是在測試完其良率跟不同參數條件之間的關係後，我們發現傳統的 SA 並不能準確地分辨此 Twin-8T 架構得出的電壓值以及其代表的 MAC 值。

為了準確讀值，我們與 IMDL 實驗室鐘彥麟學長討論出一款具有 AUTO-ZERO 功能，且在操作過程能透過正回饋提高準確率的新的 SA 架構 (如下右圖)。此新 SA 架構在 offset 僅為 10 毫伏時，在蒙地卡羅一萬個點測試結果中沒有出現任何讀取錯誤，使我們讀取 CIM 時的能力、準確率大幅上升。



Conclusion

成功的利用 Twin-8T Cell 實現 CIM，並藉由提出新的感測器 (SA) 架構幫助我們對 CIM 結果準確地進行讀值。

References

- [1] Xin Si et al., "A Twin-8T SRAM Computation-in-Memory Unit-Macro for Multibit CNN-Based AI Edge Processors", *IEEE J. Solid-State Circuits*, vol. 55, pp. 189-202, November 2019.
- [2] B. Wicht, T. Nirschl and D. Schmitt-Landsiedel, "Yield and speed optimization of a latch-type voltage sense amplifiers", *IEEE J. Solid-State*