

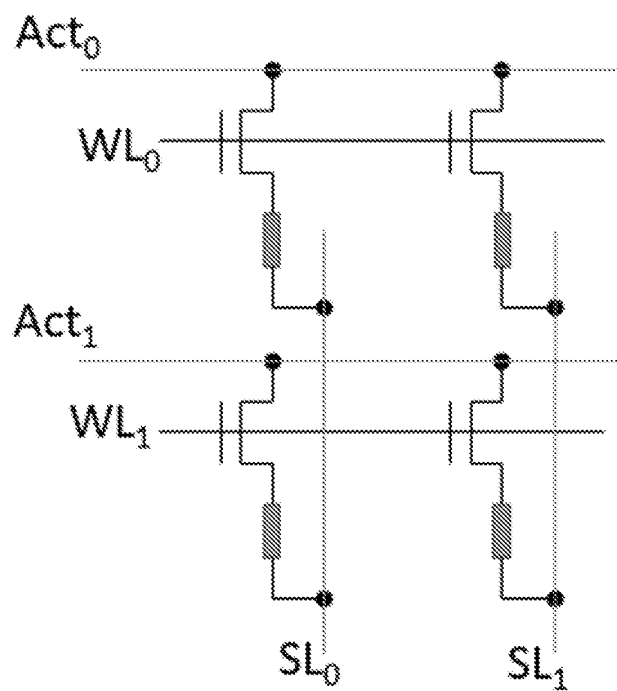


FIG. 1

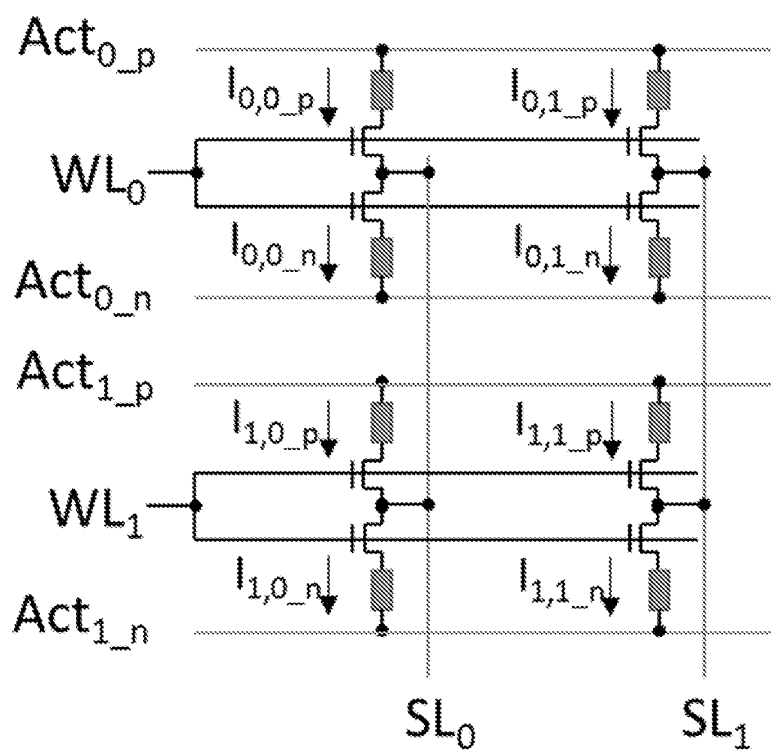


FIG. 2

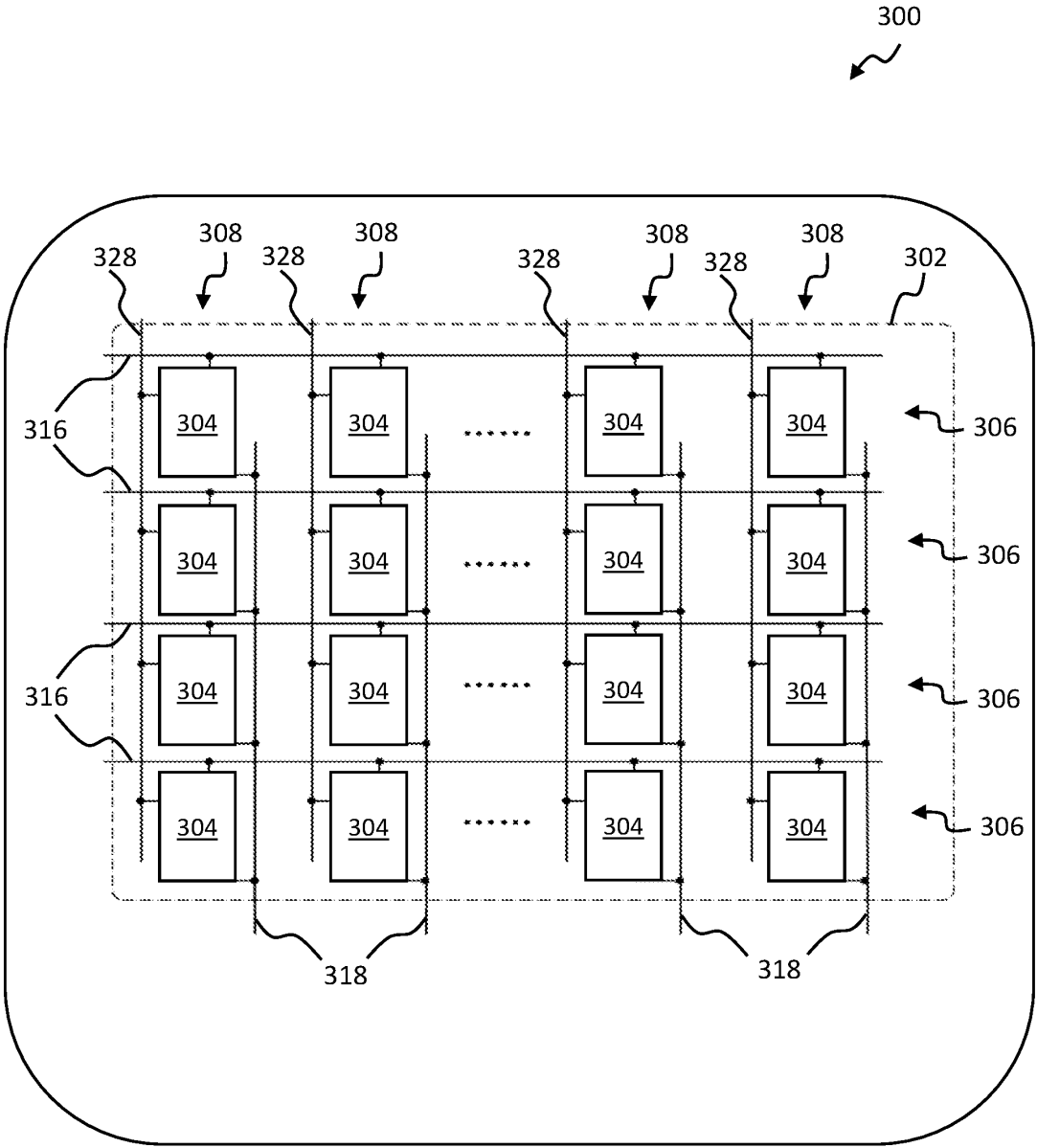


FIG. 3

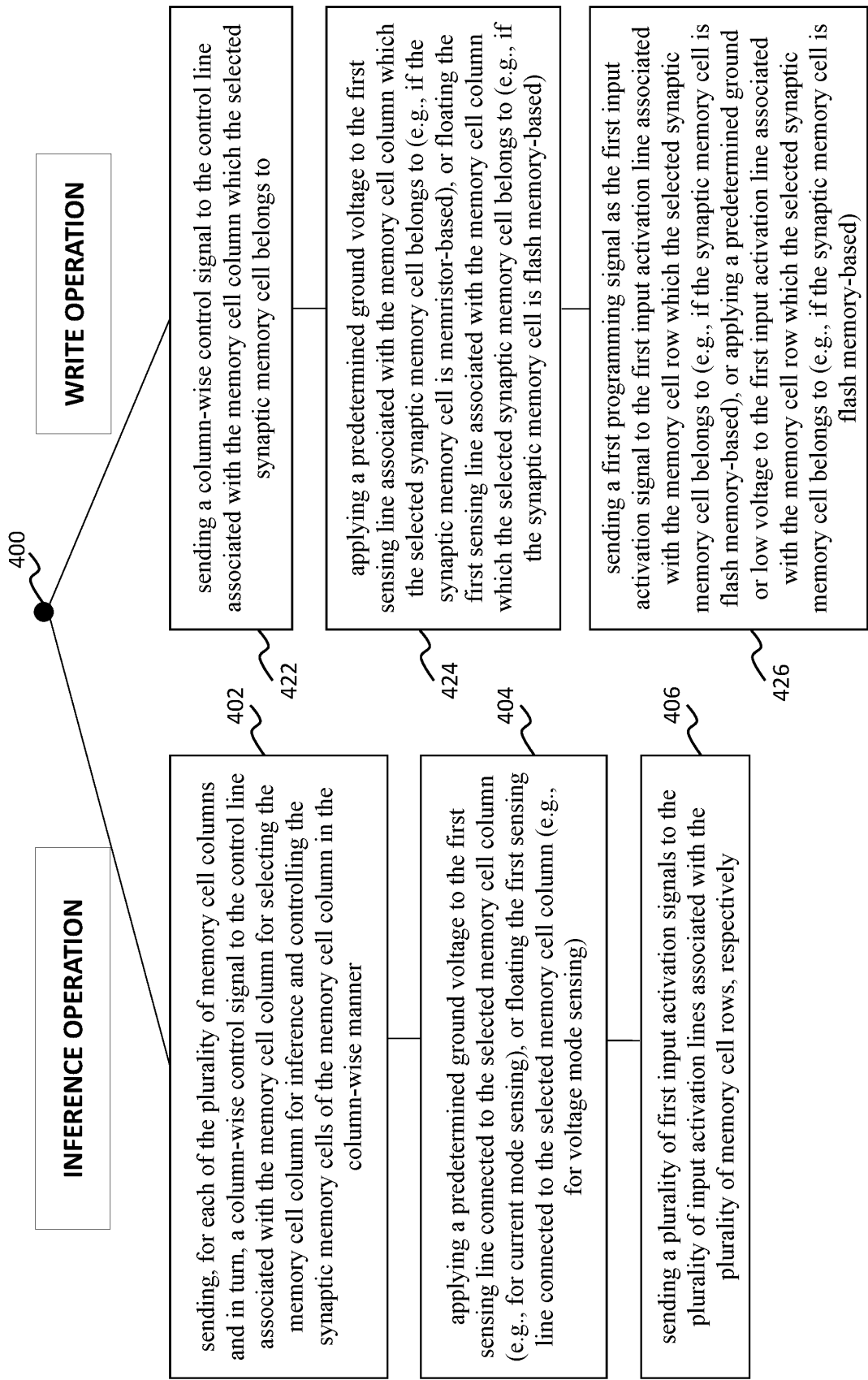


FIG. 4

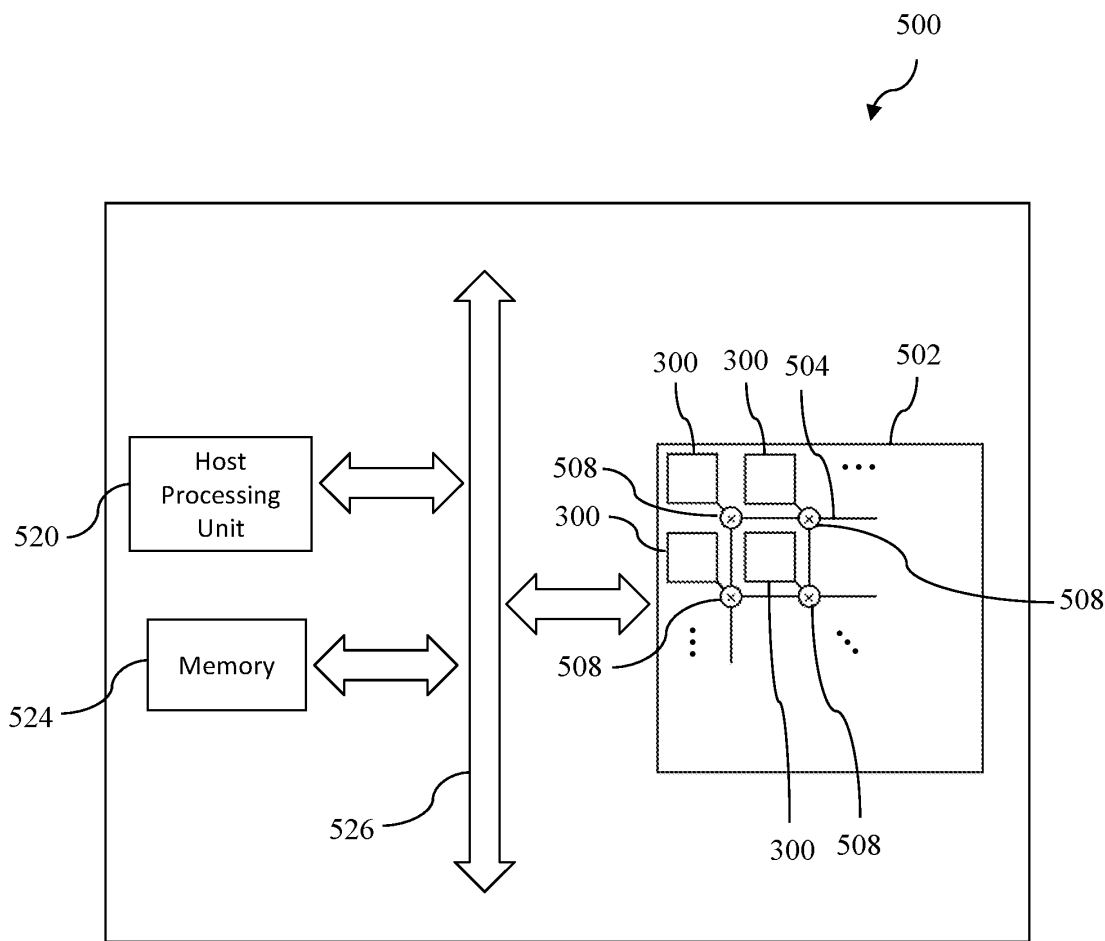


FIG. 5A

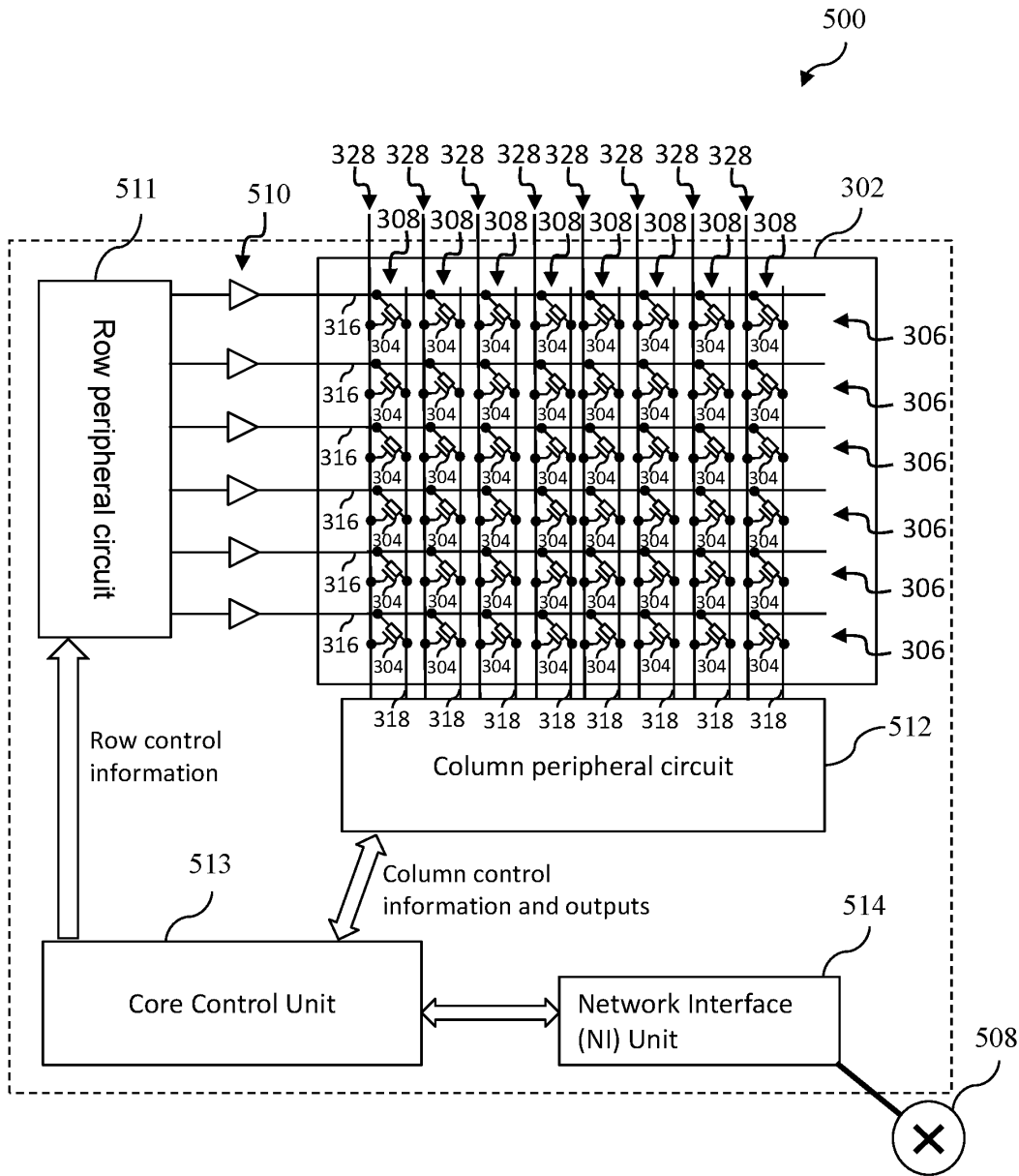


FIG. 5B

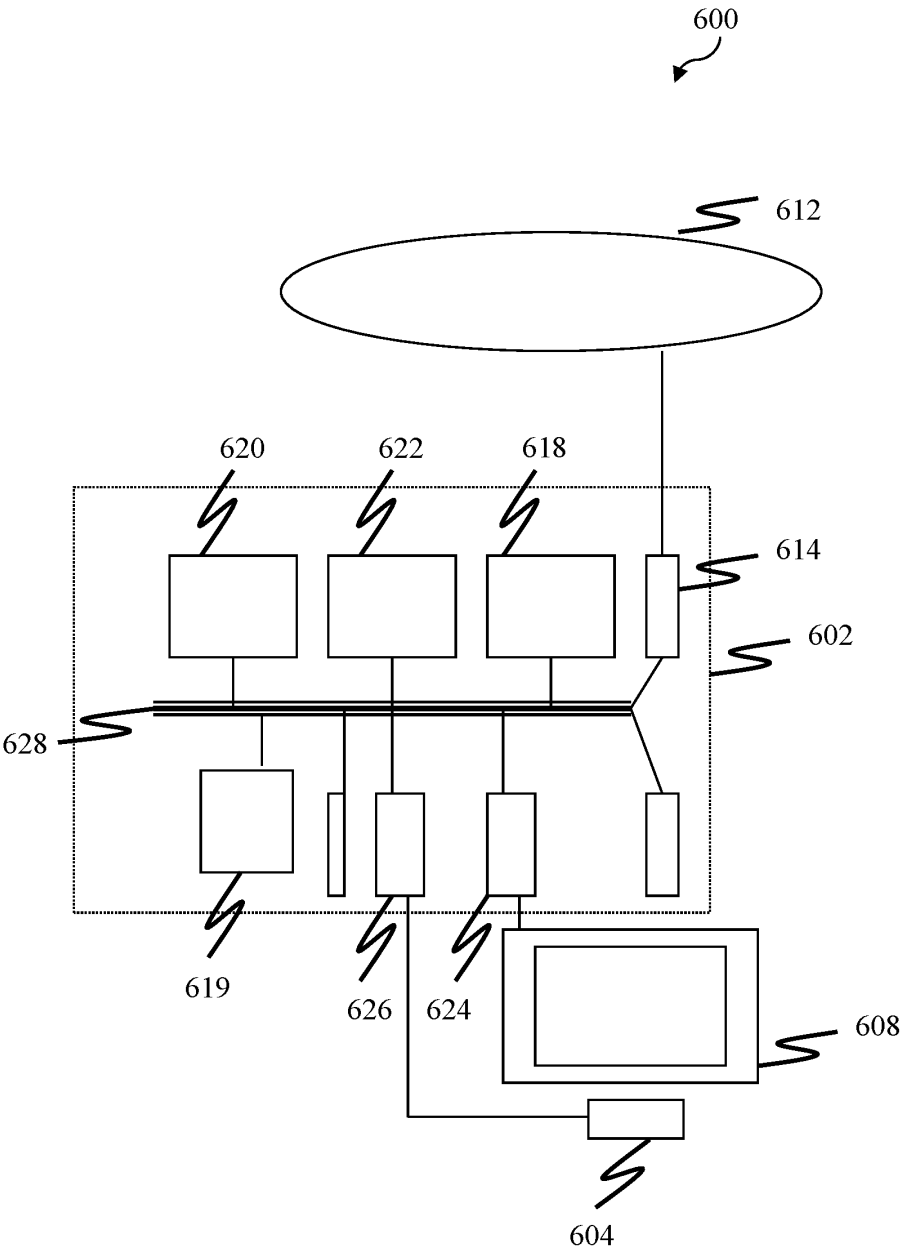


FIG. 6

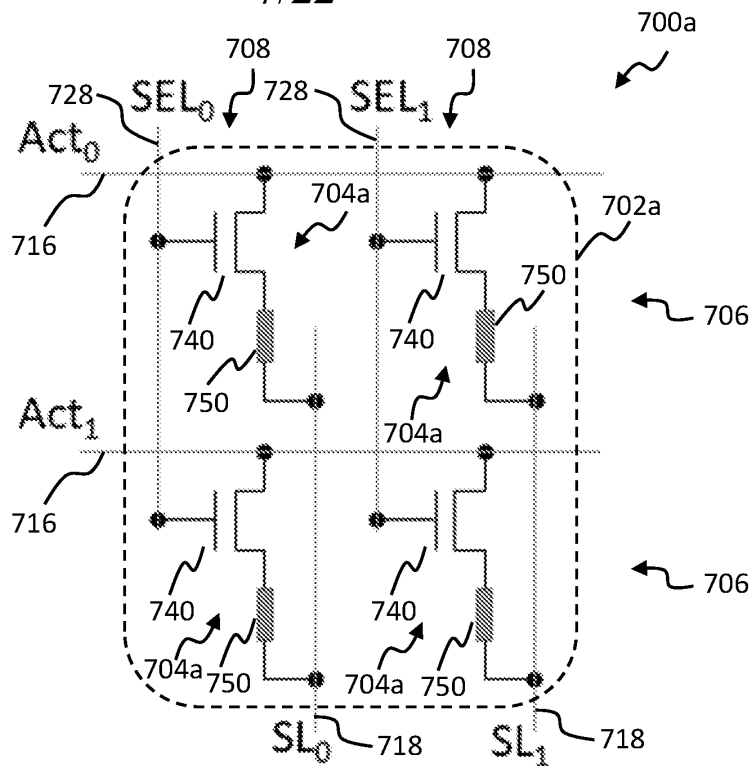
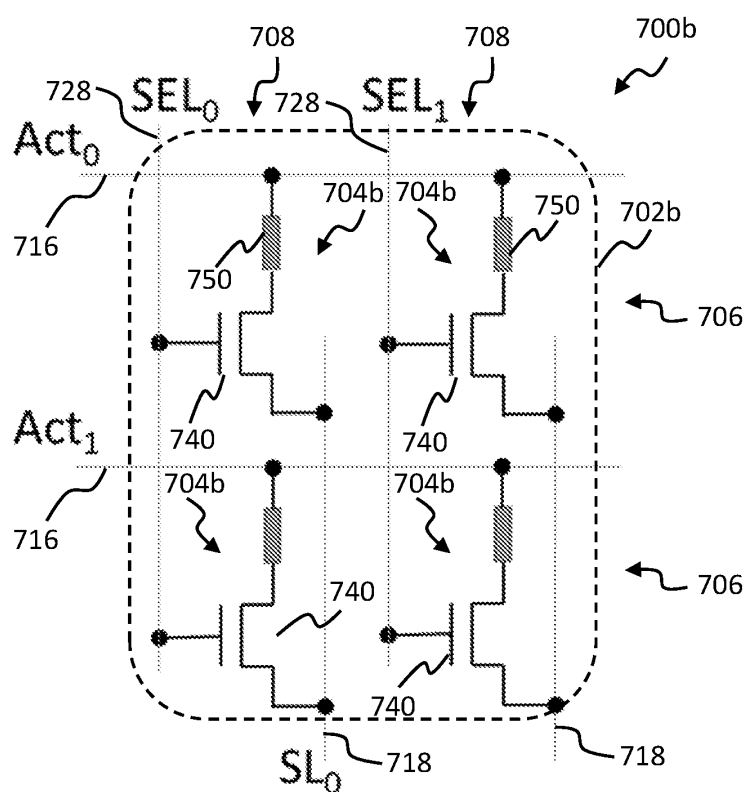


FIG. 7A



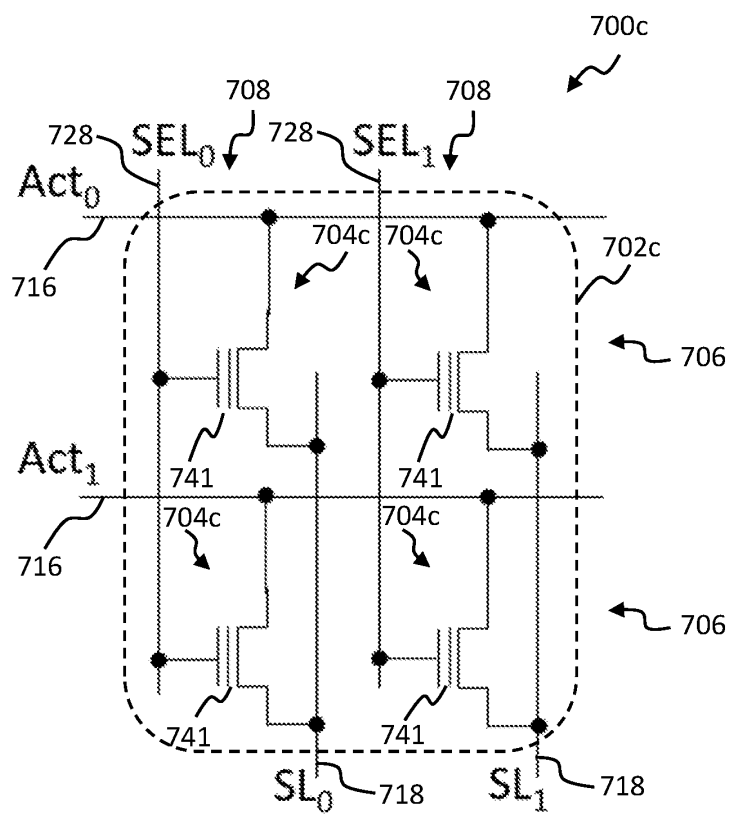


FIG. 7C

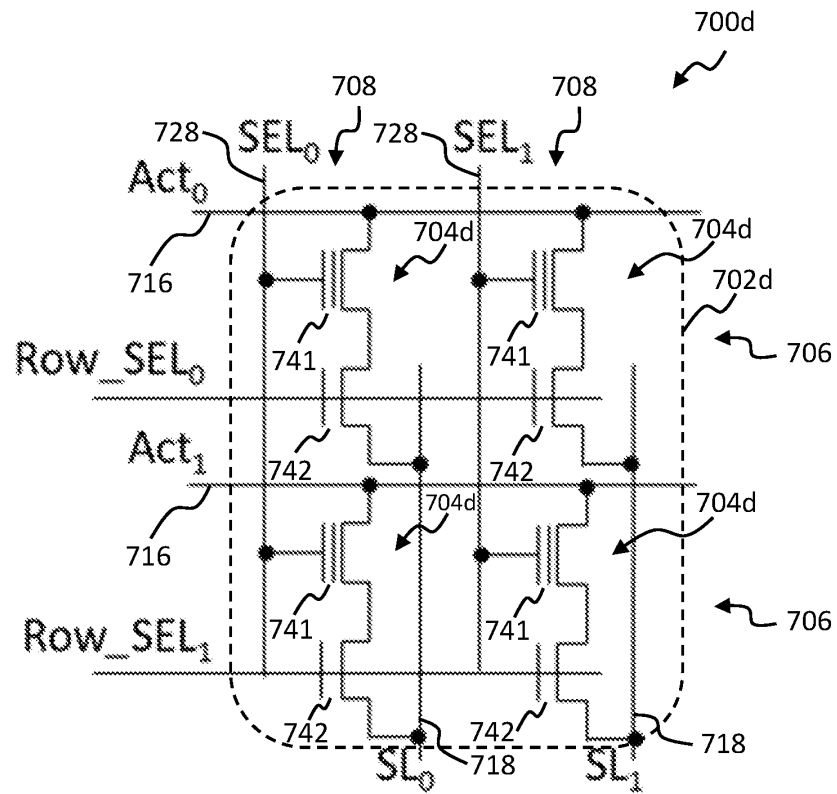


FIG. 7D

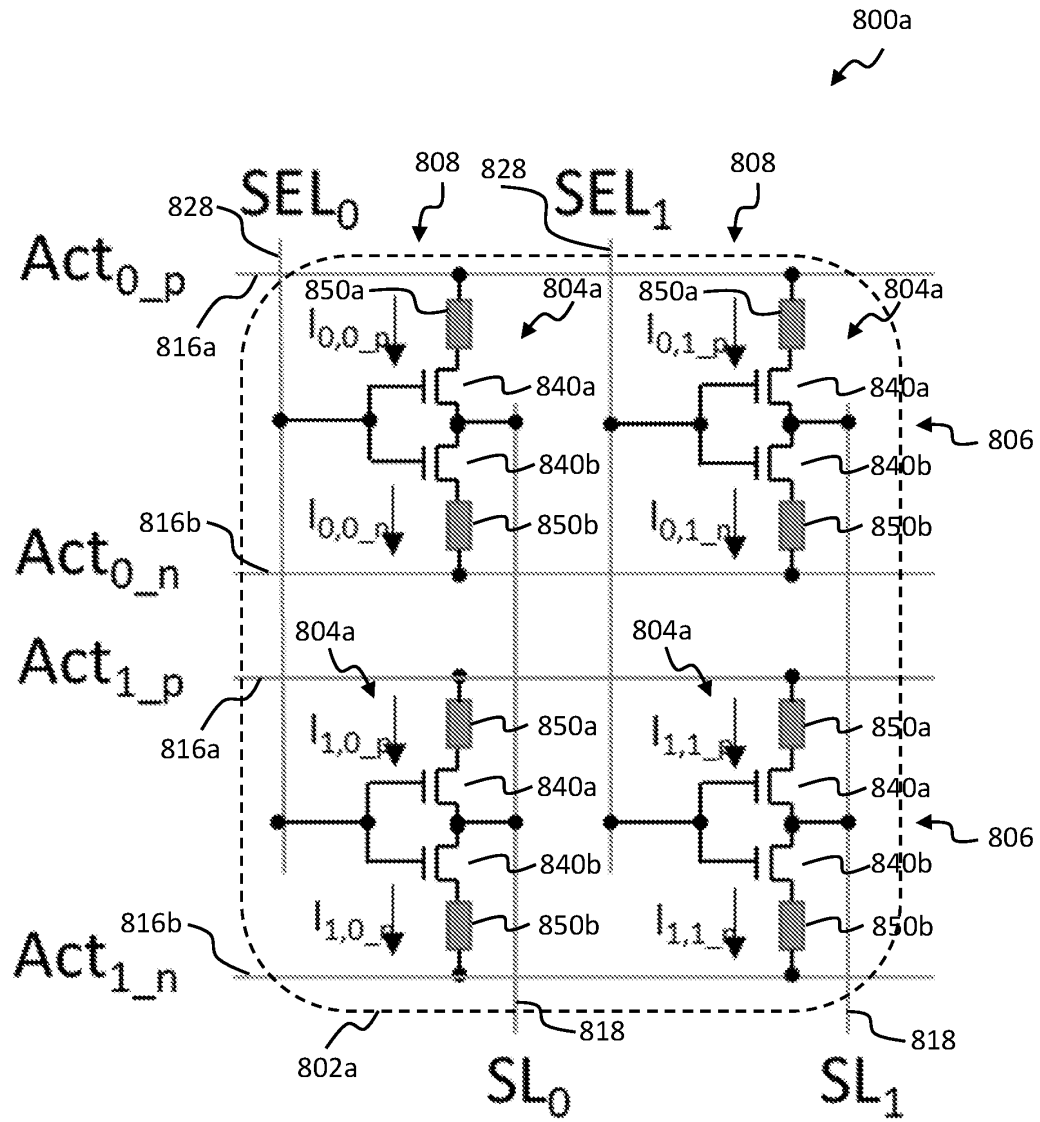


FIG. 8A

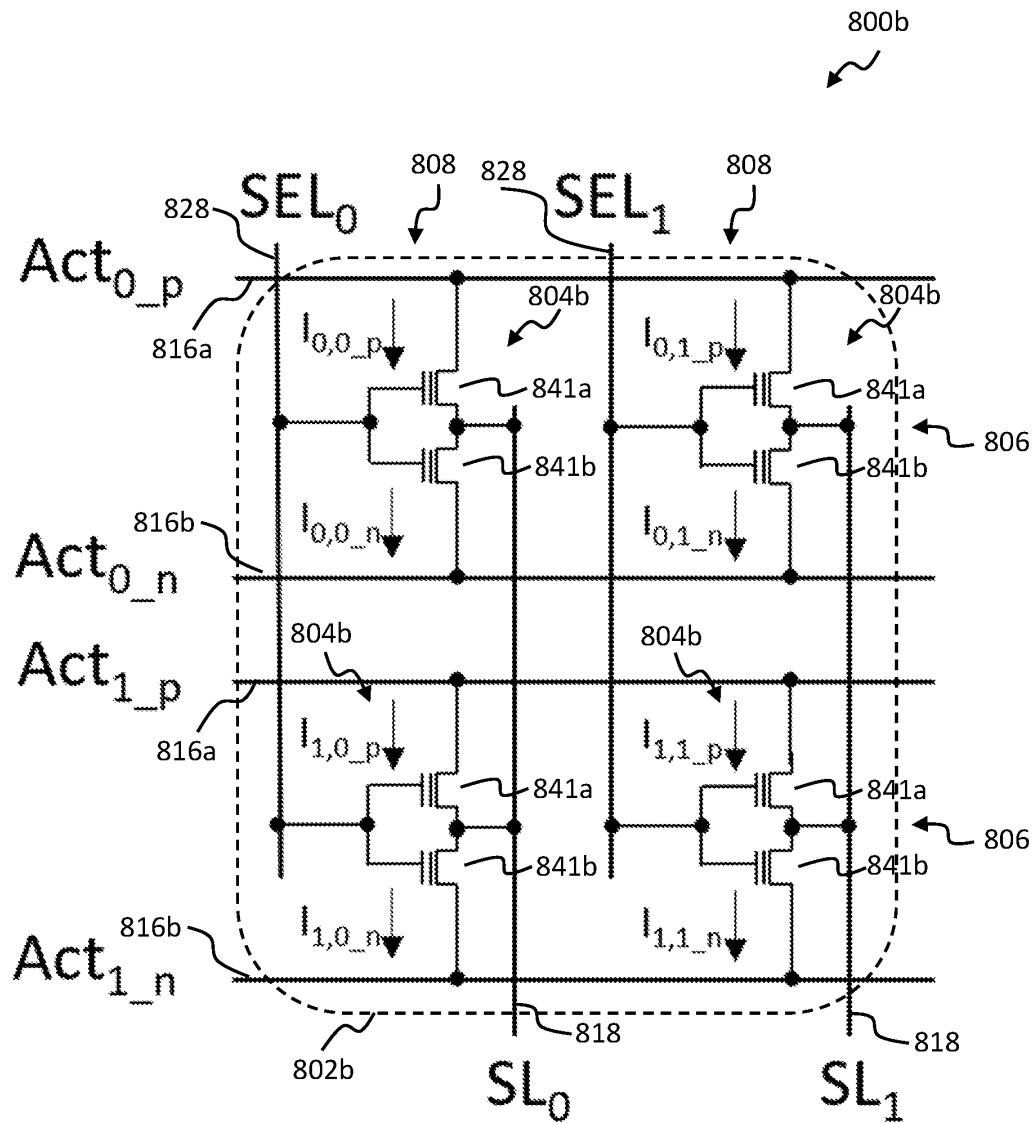


FIG. 8B

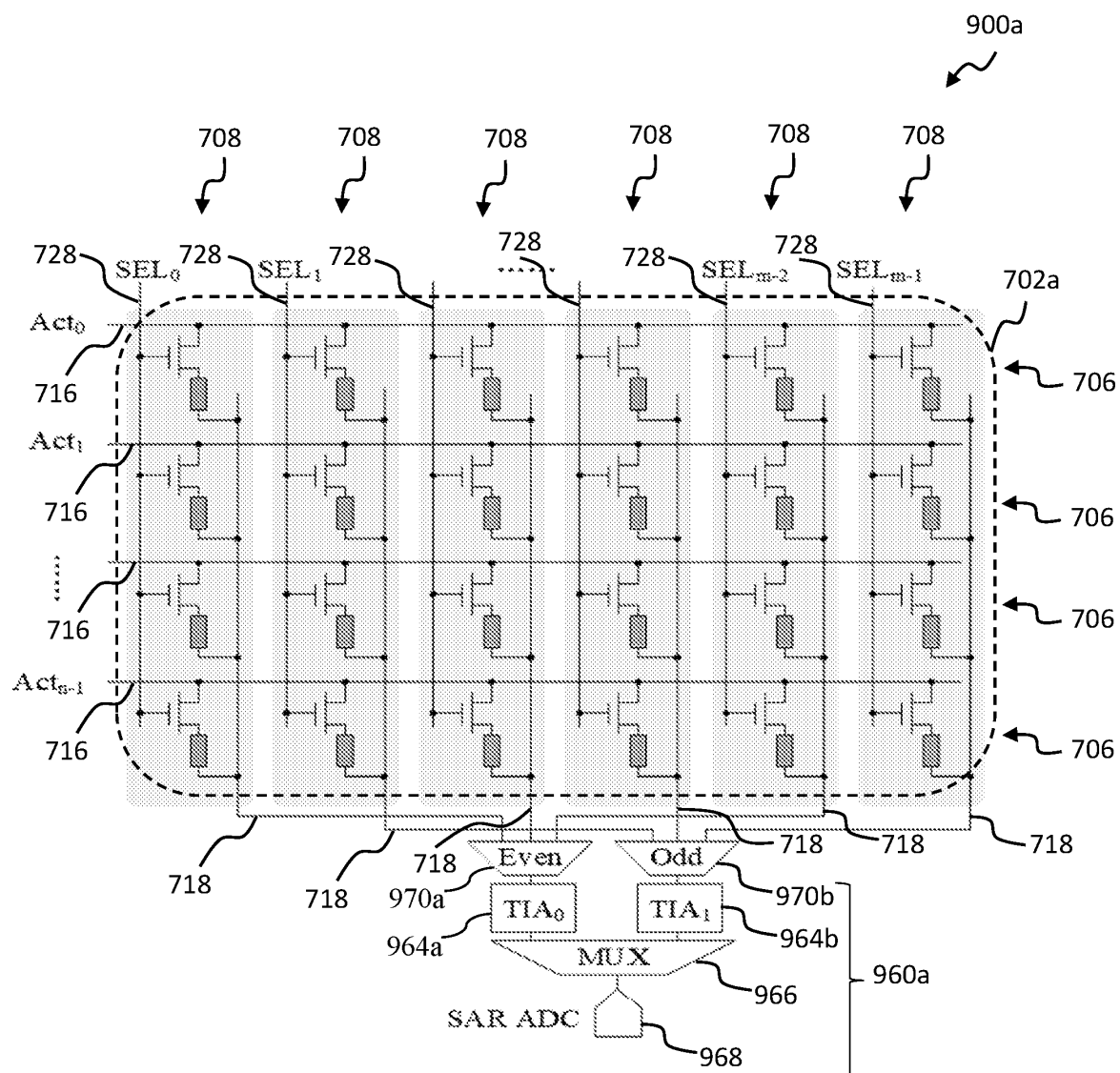


FIG. 9A

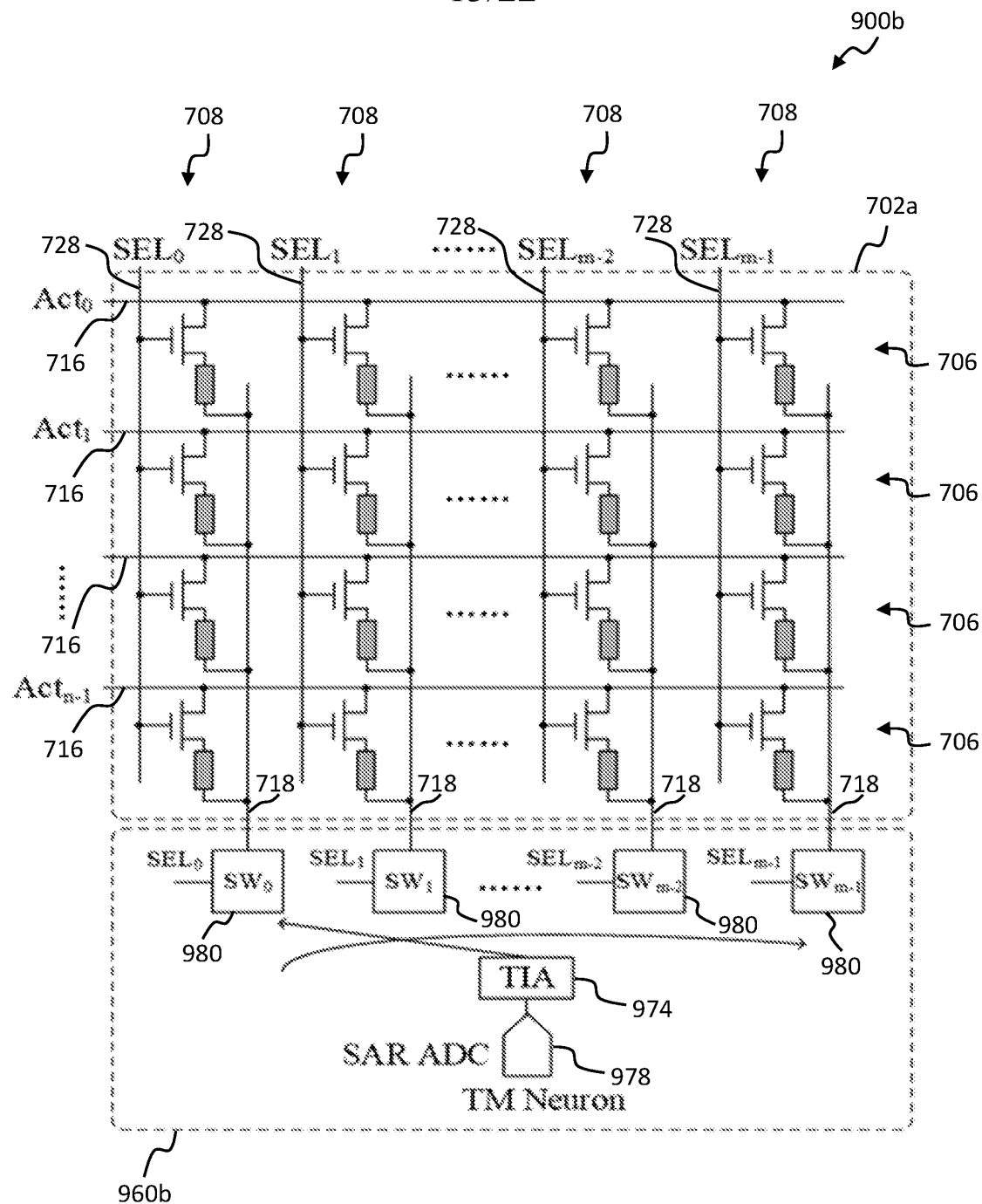


FIG. 9B

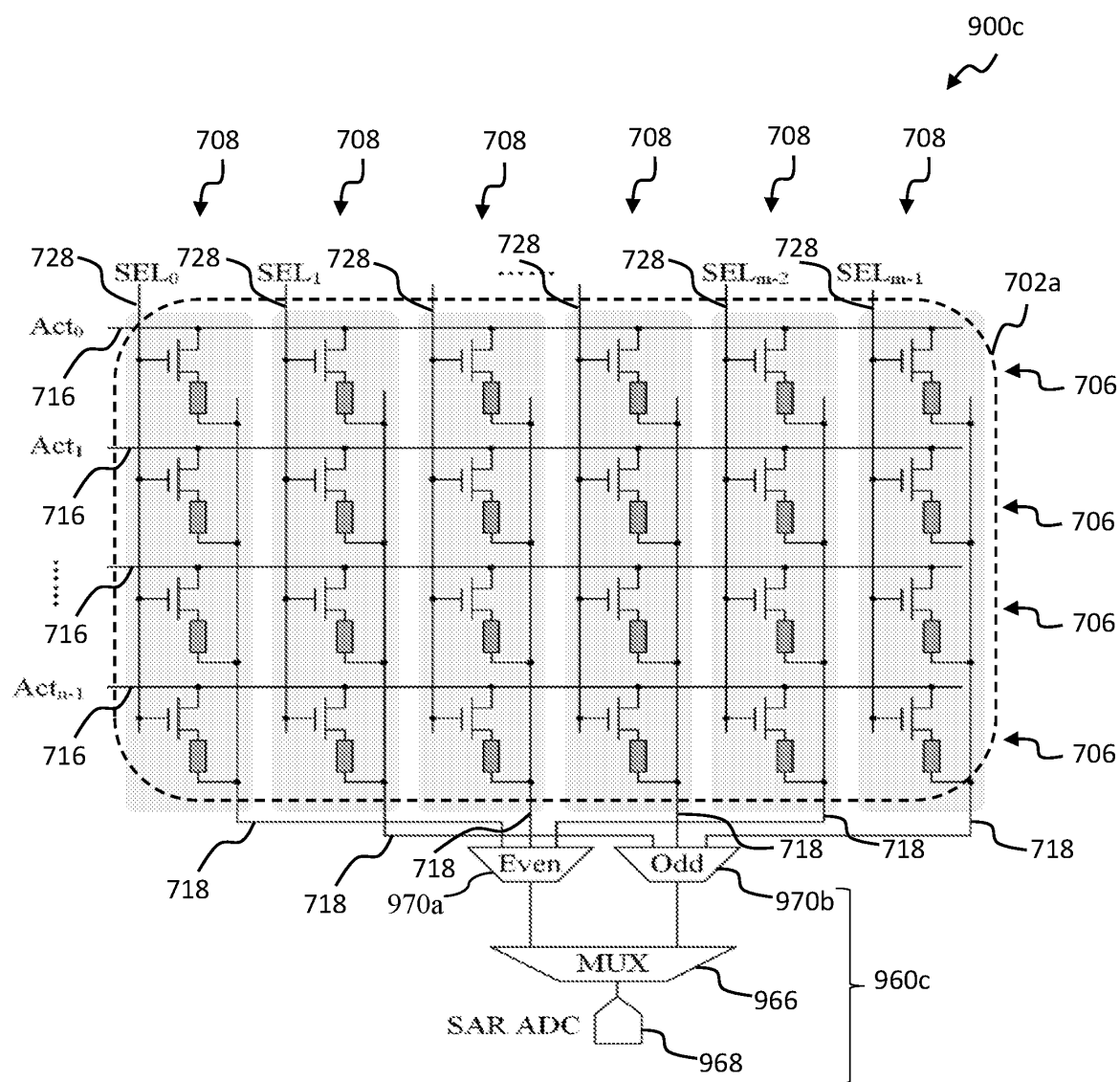


FIG. 9C

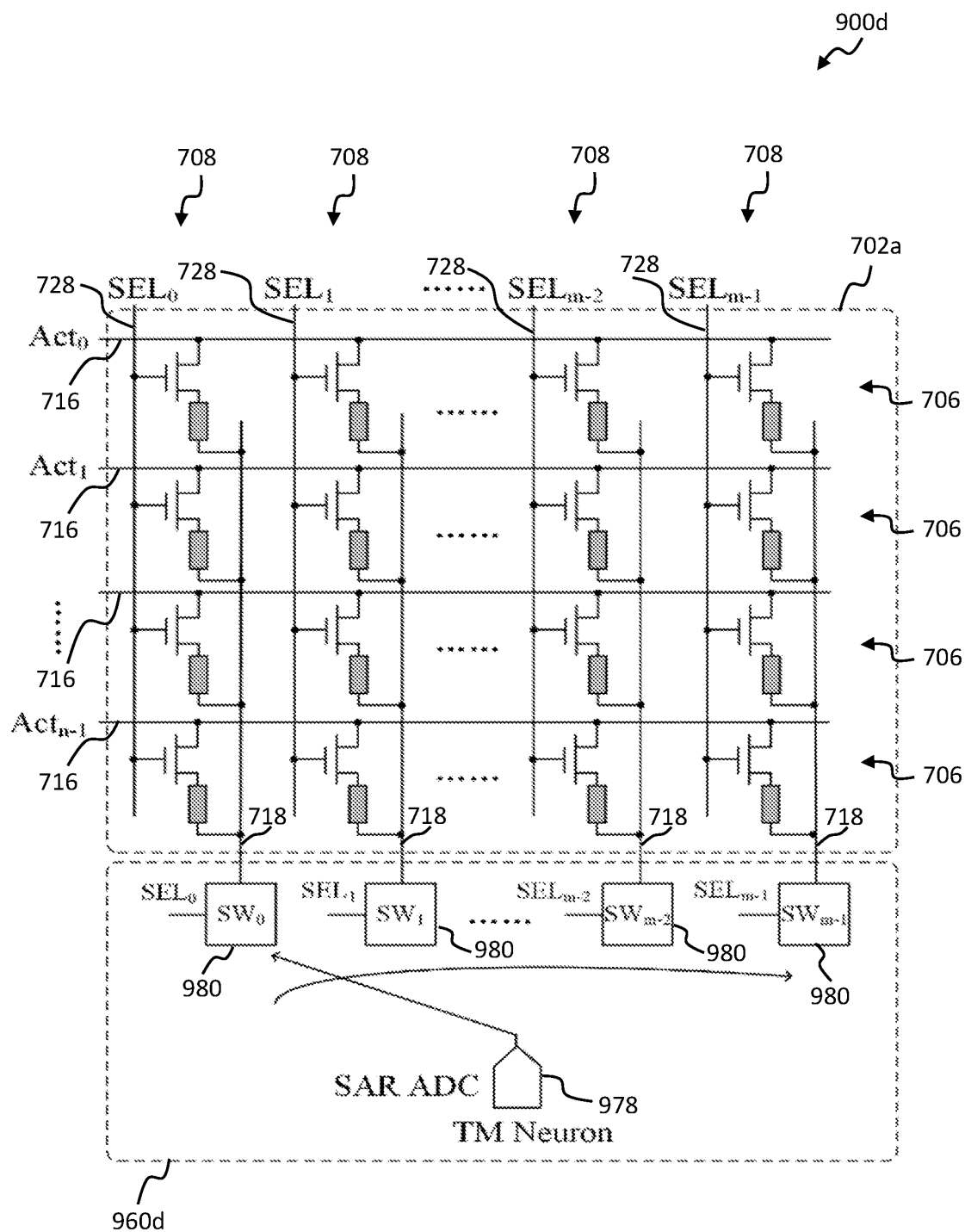


FIG. 9D

Table 1

	Component		Quantity	Area (mm ²)	Peak Power (mW)	Latency (ns)	Energy/MAC (pJ)
No-TM	Array	1T1R	256×256	0.011	65.636	10	0.010
		2T2R	256×256	0.022	65.536	10	0.010
	DAC+OP-AMP*		256	0.100	15360	10	2.343
	S/S ADC		256	0.768	51.200	200	0.156
	Total	Analog Input	-	0.879	15476.736	210	2.509
		Digital Input	-	0.890	15476.736	210	2.509
TM	Array	1T1R	-	0.779	116.736	840**	0.665**
		2T2R	-	0.790	116.736	840**	0.665**
	DAC*	1T1R	256×256	0.011	0.256	2560	0.010
		2T2R	256×256	0.022	0.256	2560	0.010
	OP-AMP*	256	0.013	0.256	10	0.010	
		256	0.003	1.280	10	0.050	
	SW	1	0.003	-	-	-	
	TIA	1	0.002	0.500	2560	0.019	
	SAR ADC	1	0.013	1.200	2560	0.047	
		Digital Input	-	0.029	1.956	10280**	0.308**
Total	2T2R	-	0.040	1.956	10280**	0.308**	
	Present	1T1R	-	0.045	3.492	5140	0.136
2T2R		-	0.056	3.492	5140	0.136	

* The DAC and OP-AMP are only used in analog-input architecture.
** The latency and energy consumption are increased due to multiple cycles.

FIG. 10

Table 2

Component	Quantity	Area (mm ²)	Peak Power (W)	Latency per Image (ms)	Energy per Image (mJ)
2T2R array	2121	46.983	0.042	32.113	0.154
DAC	542976	27.149	0.042	32.113	0.155
OP-AMP	542976	5.430	0.208	32.113	0.774
SW	2121	6.363	-	-	-
TIA	2121	4.242	0.133	32.113	0.068
SAR ADC	2121	27.573	0.318	32.113	0.817
Total	-	117.739	0.742	64.225	1.968

FIG. 11**Table 3**

Architecture		Area (mm ²)	Peak Power (W)	Latency per Image (ms)	Energy per Image (mJ)
No-TM	Analog Input	1887.997	2527.996	10.537	59.256
	Digital Input	1675.911	30.376	20.070	11.517
TM	Digital Input	85.161	0.492	128.451	4.159
	Present	117.739	0.742	64.225	1.968

FIG. 12**Table 4**

Layer	# of operations	Effective columns in array	# of array	# of ADC for each array/layer	Area (mm ²)	Peak Power (W)	Latency (ms)
Conv3-64	224×224	64	1	32/32	0.693	0.060	2.007
Conv3-64			3	32/96	2.079	0.274	
Conv3-128	112×112	128	3	16/48	1.092	0.137	2.007
Conv3-128			5	16/80	1.820	0.248	
Conv3-256	56×56	256	5	8/840	0.997	0.124	2.007
Conv3-256			9	8/72	1.795	0.235	
Conv3-256			9	8/72	1.795	0.235	
Conv3-512	28×28	256	18	2/36	1.369	0.091	2.007
Conv3-512			36	2/72	2.739	0.182	
Conv3-512			36	2/72	2.739	0.182	
Conv3-512	14×14	256	36	1/36	1.998	0.093	1.004
Conv3-512			36	1/36	1.998	0.093	
Conv3-512			36	1/36	1.998	0.093	
FC4096	512×7×7×4096	256	1568	1/1568	87.042	0.112	0.253
FC4096	4096×4096	256	256	1/256	14.211	0.112	0.253
FC4096	4096×1000	256	64	1/64	3.553	0.112	0.253
Total	-	-	2121	-/2616	127.916	2.162	2.007

FIG. 13

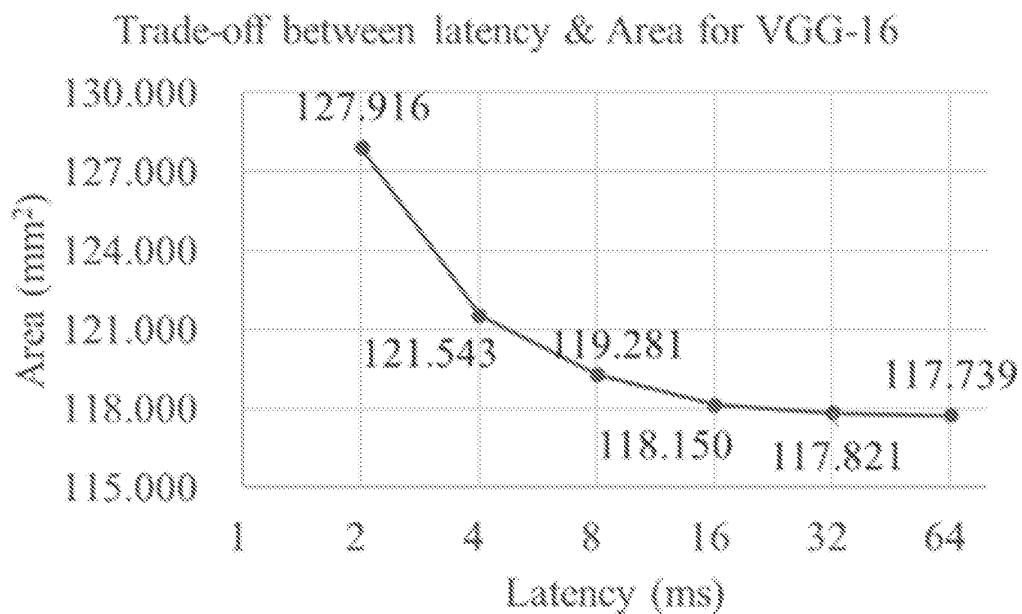


FIG. 14A

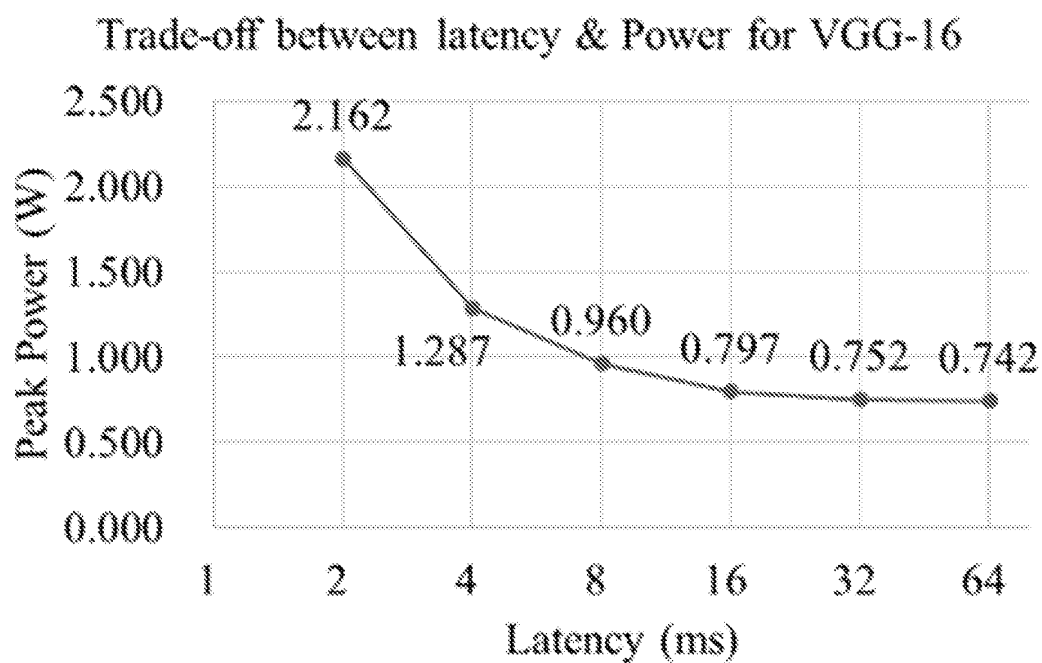


FIG. 14B

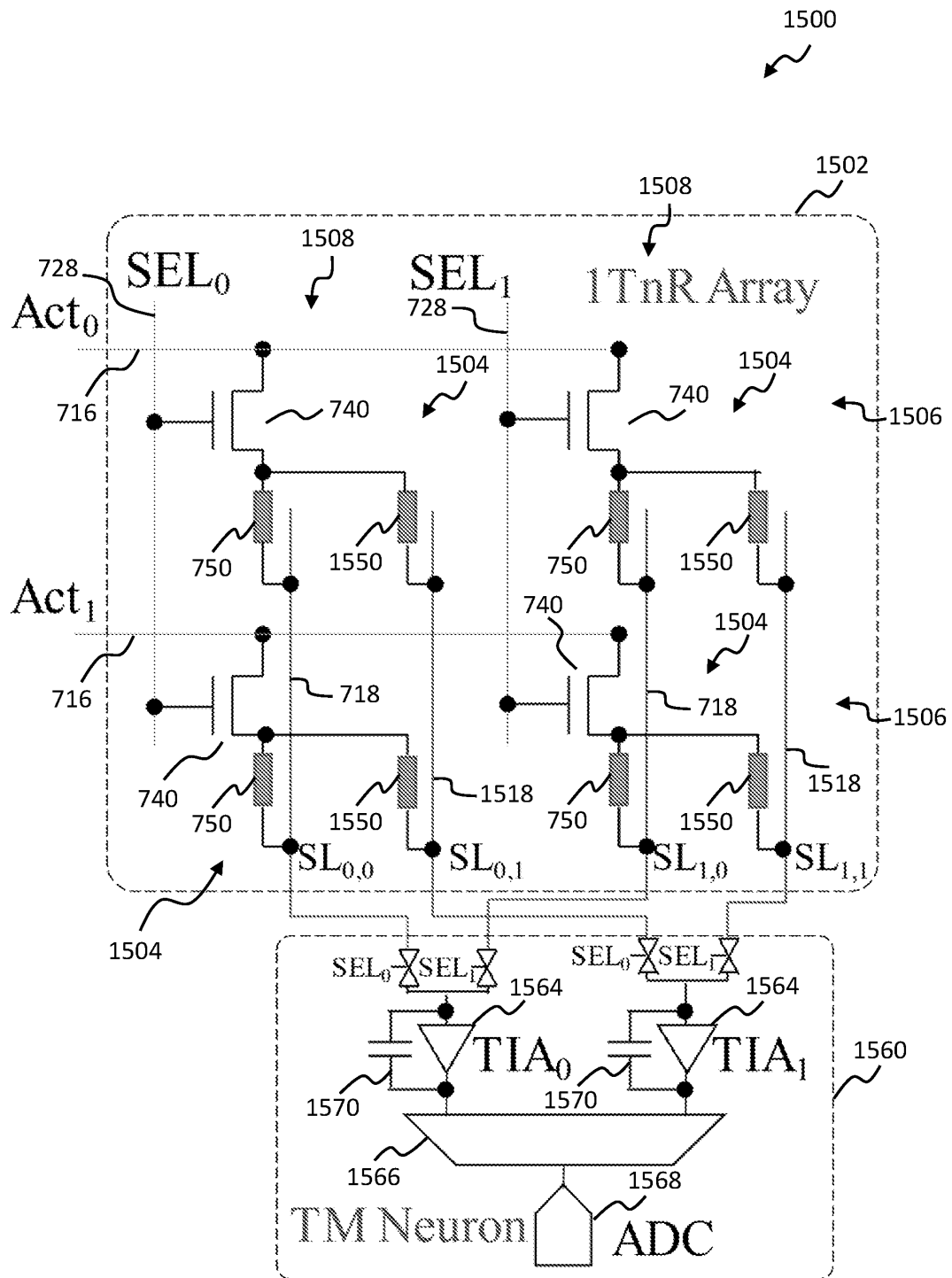


FIG. 15

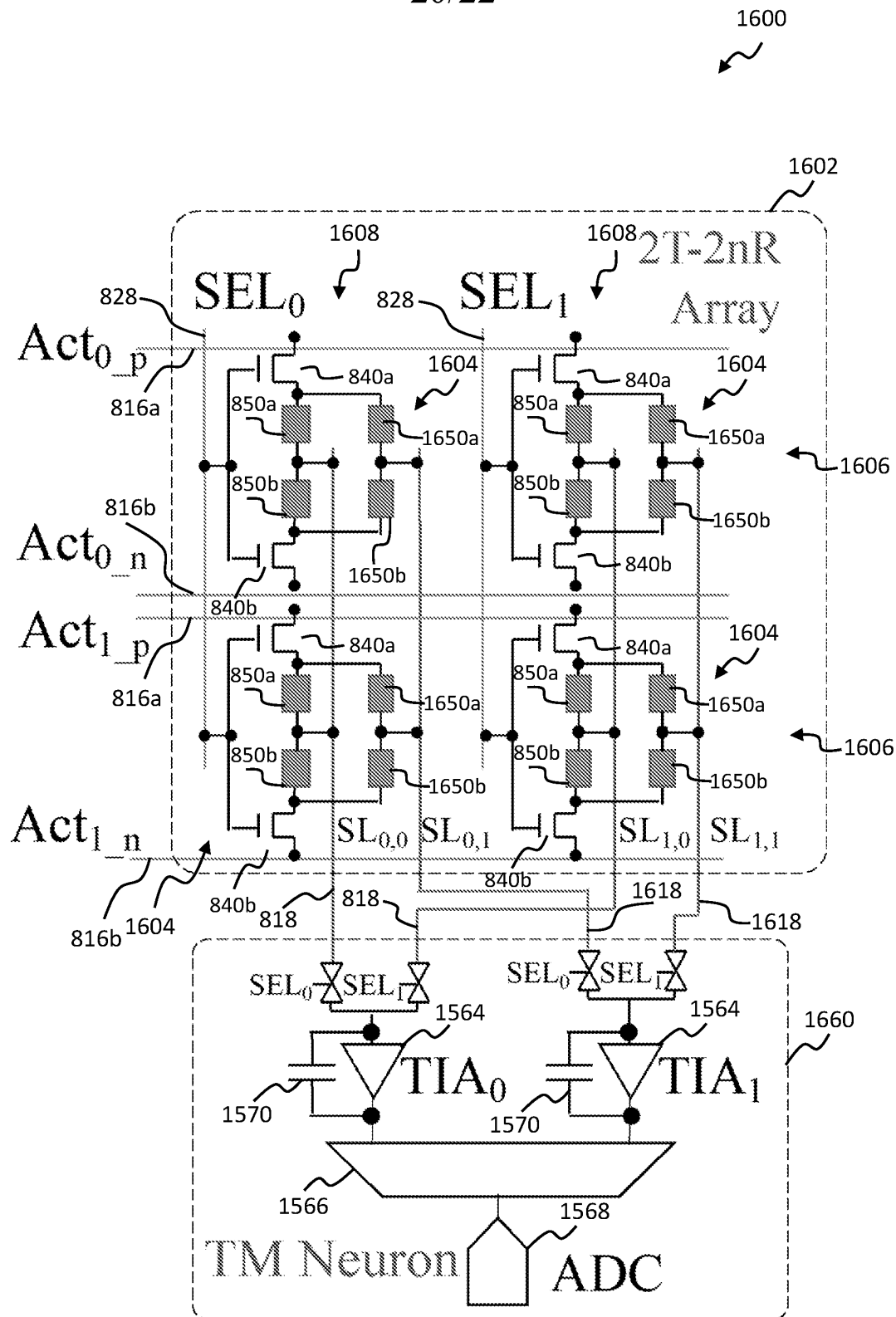


FIG. 16

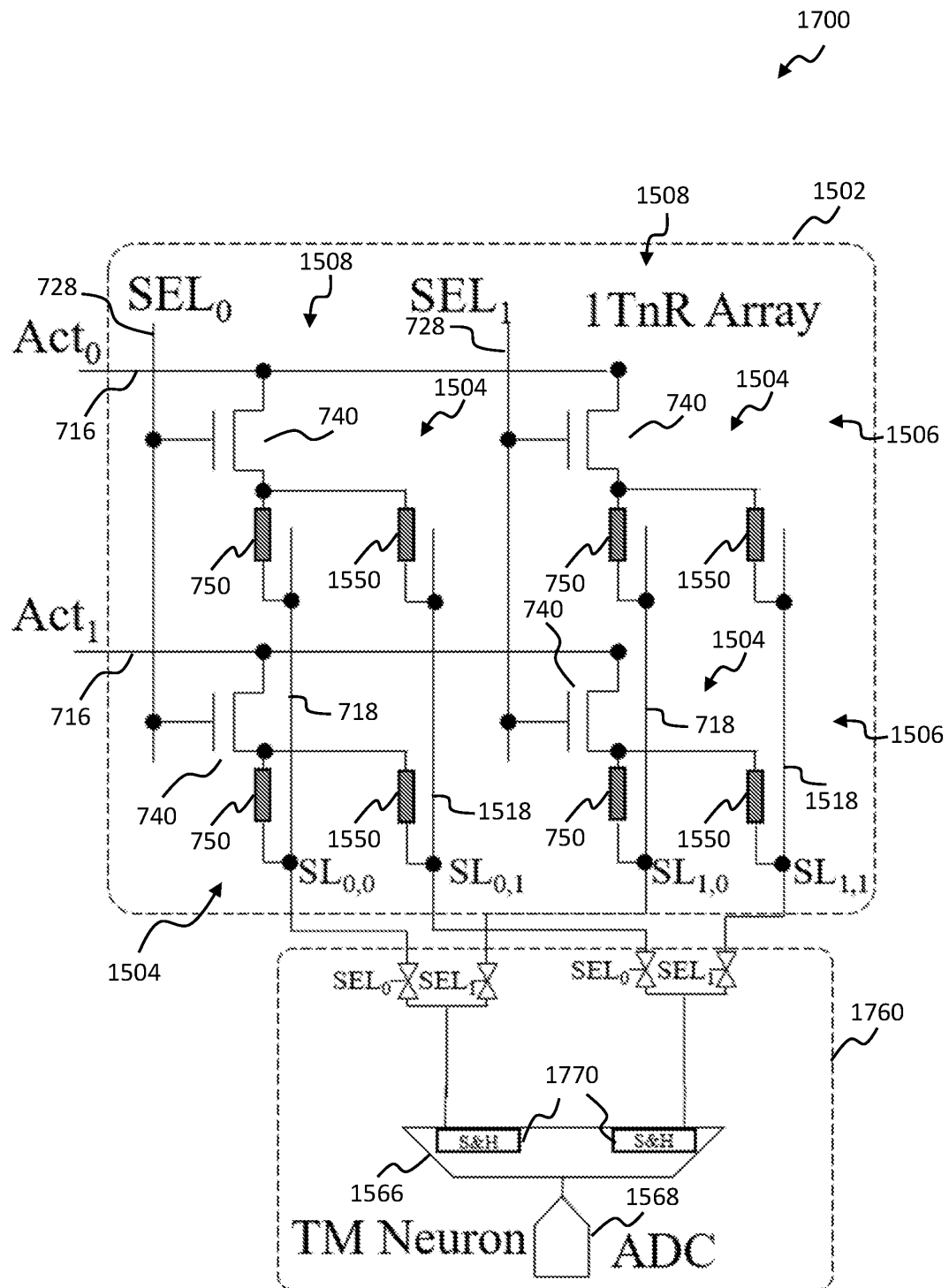


FIG. 17

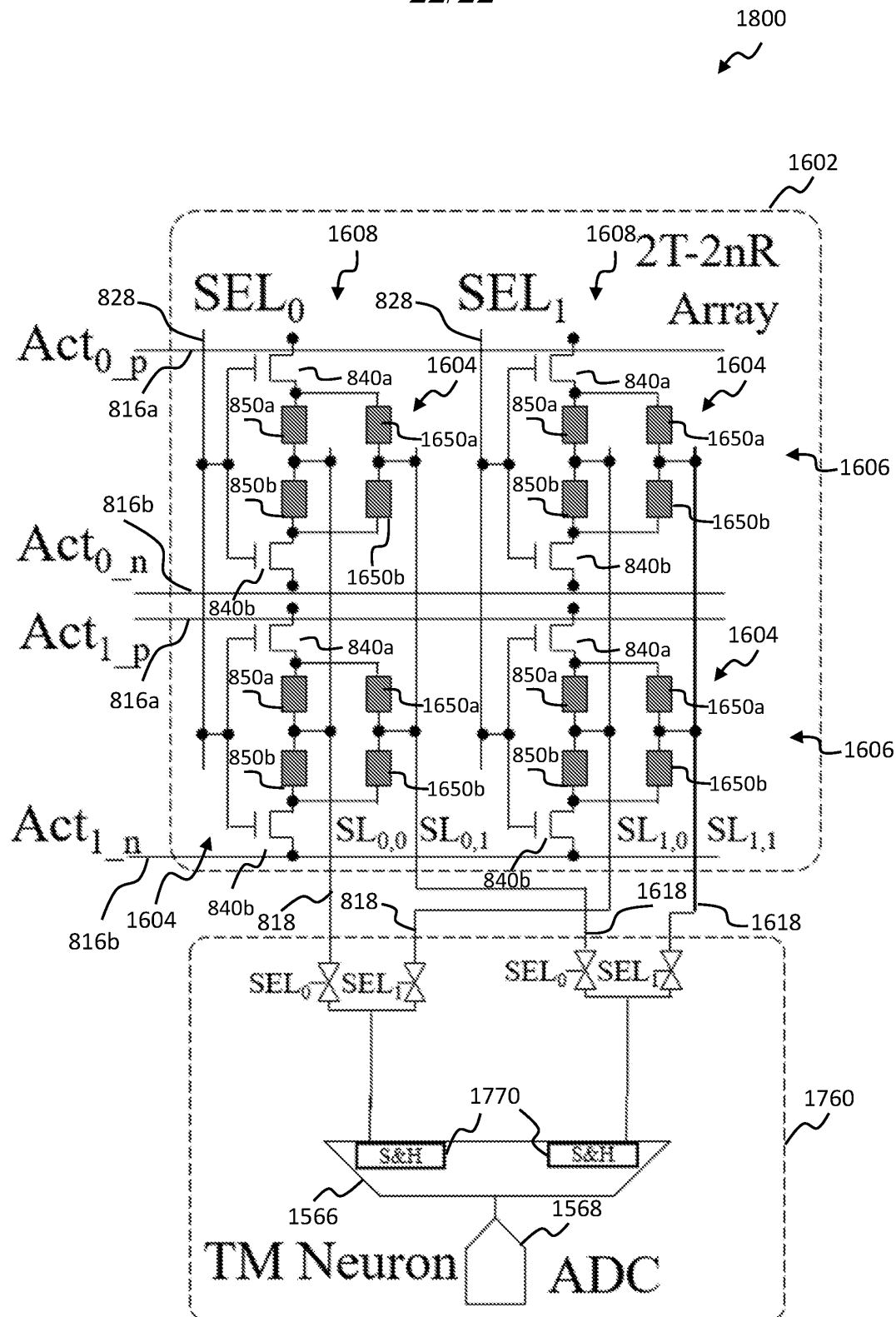


FIG. 18