



#### 공학석사 학위논문

# OTS switching delay를 활용한 RRAM 기반 발화 지연 뉴로모픽 시냅스 소자 개발

2023년 2월

서울대학교 대학원

재료공학부 김영훈

## OTS switching delay를 활용한 RRAM 기반 발화 지연 뉴로모픽 시냅스 소자 개발

#### 지도 교수 김 상 범

이 논문을 공학석사 학위논문으로 제출함 2023년 01월

> 서울대학교 대학원 재료공학부 김 영 훈

김영훈의 공학석사 학위논문을 인준함 2023년 01월

위원장_	박 민 혁	(인)
부위원장 _	김 상 범	(인)
_ 위 원_	강 기 훈	(인)

#### 초 록

인공지능의 발전에 따라 보다 효율적인 연산을 위해 실제 생물 신경망을 모사한 회로의 개발에 연구가 집중되고 있다. 실제 뉴런과 시냅스의 모습을 본 딴 crossbar array 구조가 많이 이용되고 있지만 해당 구조는 실제 뉴런과 다르게 뉴런마다 다른 delay를 줄 수 없다는 한계를 가지고 있다. 본 연구에서는 이러한 한계점에 착안하여 뉴런 별로 다른 delay를 부여할 수 있는 새로운 시냅스 회로 구조를 제시하였다. 새로운 시냅스 회로 구조에서는 Ovonic Threshold Switch(OTS)와 Resistive random access memory(RRAM)을 직렬 연결하여 RRAM의 가변저항을 바꾸어 OTS에 인가되는 전압을 바꿈으로써 각 뉴런마다 다른 delav를 부여할 수 있다. OTS와 RRAM을 통합한 소자를 만들기 위해 RRAM과 OTS 단일 소자의 특성 최적화를 진행했다. RRAM은 Switching layer 두께와 compliance current를 조절한 결과 기존 동작 확률 10%, endurance는 94 cycles에서 동작 확률 100%, endurance 1000 cycles 이상으로 크게 성능을 개선할 수 있었다. OTS 소자는 기존 GeSe계 switching layer를 AgGeSe로 바꾼 결과 기존 1-10 us의 delay에서 4.8 ms까지 늘릴 수 있었다. 이는 실제 뉴런의 발화 지연에 상당하는 수치이다. RRAM과 OTS를 통합한 소자를 제작 완료했지만 소자 동작 특성은 확인할 수 없었다. 이를 개선하기 위해서는 etch rate 등 공정의 최적화를 더 진행해야 한다고 판단된다.

주요어 : 오보닉 문턱 스위치, 저항 변화 메모리, 저항 스위칭, 임계 스위칭, 스위칭 지연, 뉴로모픽 시냅스 학 번 : 2021-22969

iii

1.		서 론	1
	1.1.	발화지연 뉴로모픽 시냅스 소자의 필요성	1
2.		발화 지연 시냅스 소자 구조 설명	5
	2.1.	RRAM 소자 동작 특성	5
	2.2.	OTS 소자 동작 특성	7
	2.3.	발화 지연 시냅스 소자 동작 특성	8
3.		실험 및 분석 방법	10
	3.1.	RRAM 소자 제작 및 측정	10
	3.1.1.	RRAM 소자 제작	10
	3.1.2.	RRAM 소자 측정	13
	3.2.	OTS 소자 제작 및 측정	15
	3.2.1.	OTS 소자 제작	15
	3.2.2.	OTS 소자 측정	16
	3.3.	RRAM+OTS 소자 제작 및 측정	17
	3.3.1.	RRAM+OTS 소자 제작	17
	3.3.2.	RRAM+OTS 소자 측정	20
4.		결과 및 논의	21
	4.1.	RRAM 소자 제작 및 측정	21
	4.1.1.	RRAM 단일 소자 switching 성능 개선 실험	21
	4.1.2.	RRAM 단일 소자 pulse test	29
	4.2.	OTS 소자 제작 및 측정	31
	4.2.1.	N-doped GeSe OTS 소자 제작 및 측정	31
	4.2.2.	N-doped AgGeSe OTS 소자 제작 및 측정	34
	4.3.1.	RRAM+OTS 소자 제작	39
	4.3.2.	RRAM+OTS 소자 측정	41
4.	결론		42
참	고 문헌		43
Ał	ostract		47

## 표 목차

Table 3.1.1.1 RRAM에 사용된 물질, 증착 장비, 박막 두께......12 

그림 목차

Figure 1.1.1 Crossbar array 구조......2 Figure 1.1.2 기존 뉴로모픽 시스템의 회로 구조(좌)와 시냅스 구조(우) Figure 1.1.3 OTS 기반 발화 지연 시스템 회로 구조(좌)와 시냅스 구 Figure 2.3.1 OTS 기반 발화 지연 시냅스 소자 구조(좌)와 서로 다른 가변저항 값에 따라 변화하는 delay 모식도(우) ......8 Figure 3.1.1.1 RRAM 단일 소자 구조 ......10 Figure 3.1.1.2 RRAM 단일 소자 제작에 사용된 쉐도우 마스크 도면11 Figure 3.1.2.1 RRAM DC I-V sweep 측정 방식......13 Figure 3.2.1.1 N-doped GeSe계 OTS 소자 모식도(좌)와 소자 단면 SEM(우).....15 Figure 3.3.1.1 RRAM+OTS 소자 마스크 도면 일부......17 Figure 3.3.1.2 RRAM+ OTS 소자 마스크 도면 전체......18 Figure 3.3.1.3 RRAM+OTS 소자 process flow ......19 Figure 4.1.1.2 HfO<sub>2</sub> 두께 증가 후 측정 결과 (10 nm → 15 nm) ......26 Figure 4.1.1.3 HfO<sub>2</sub> 두께 감소 후 endurance 측정 결과 (10 nm → 7 Figure 4.1.1.4 HfO<sub>2</sub> 두께 감소 후 yield test (10 nm → 7 nm).......27 Figure 4.1.2.1 RRAM switching cycle에 따른 SET, RESET 저항 값 Figure 4.2.1.2 Square pulse 인가전압 변화에 따른 delay 변화 측정 Figure 4.2.1.3 Square pulse 인가전압 변화에 따른 delay 변화 측정 Figure 4.2.2.2 N-doped AgGeSe SEM 단면과 EDS 성분 분석 .......35 Figure 4.2.2.3 AgGeSe 임계 스위칭 소자와 GeSe 스위칭 소자의 Figure 4.3.1.1 RRAM+ OTS 소자 제작 일정오류! 책갈피가 정의되어 있지 않습니다.

## 1. 서 론

#### 1.1. 발화지연 뉴로모픽 시냅스 소자의 필요성

DNN의 전례 없는 성공에 따라 인공지능이 산업과 학계를 넘어 일상 곳곳에서 쓰이기 시작했다. (Byun et al., 2022) 특히 인공지능이 이미지 인식, 자연어 처리와 같은 복잡한 작업을 처리할 수 있게 되며 자율주행, 인공지능 비서, 사물인터넷 등 인공지능이 활용되는 분야는 점차 확대되고 있다. (He, Zhang, Ren, & Sun, 2016; Shoeybi et al., 2019) 한편 1998년의 알고리즘 LeNet-5(LeCun, Bottou, Bengio, & Haffner, 1998)는 10<sup>6</sup>개 이하의 파라미터를 이용한 반면 최근 인공지능은 10<sup>14</sup>개 이상의 파라미터를 이용한다.(X. Xu et al., 2018) 이렇듯 인공지능에 사용되는 파라미터의 수가 기하급수적으로 증가하며 학습에 필요한 소비 전력 또한 기하급수적으로 증가하였다.

소비 전력의 기하급수적인 증가의 원인은 현대 컴퓨터의 폰 노이만 컴퓨팅 구조에서 찾을 수 있다. (Indiveri & Liu, 2015) 폰 노이만 컴퓨팅 구조는 메모리와 연산 장치가 분리되어 있다. 한편 인공지능 연산의 대부분은 행렬 벡터 곱 연산인데 이를 위해서는 메모리와 연산 장치 사이에 잦은 데이터 이동이 필요하다. 그런데 메모리의 동작 속도는 연산 장치에 비해 느리기 때문에 병목현상이 일어나게 된다. 이를 폰 노이만 병목현상이라고 한다.

폰 노이만 병목현상을 해결하기 위해 제시된 방법이 뉴로모픽 컴퓨팅이다. 인간 두뇌 뉴런은 높은 복잡도의 문제를 오로지 20 W만의 전력으로 해결할 수 있다. 뉴로모픽 컴퓨팅은 뉴런의 이러한 높은 효율성을 모사하기 위하여 실제 뉴런과 시냅스의 기능과 특성을 모방하는 방식으로 발전해왔다.

1



Figure 1.1.1 Crossbar array 구조

기존 뉴로모픽 컴퓨팅 시스템은 일반적으로 Figure 1.1.1와 같은 crossbar array 구조를 가진다. Crossbar array에서 각 행과 열은 뉴런에 해당하고 행과 열이 만나는 교차점은 시냅스에 해당한다. 일반적으로 이 시냅스에 가변저항 비휘발성 메모리를 위치하고 시냅스의 가중치를 가변저항의 전기전도도로 저장한다. 그 후 각 행의 뉴런에 전기 신호를 주입하면 옴의 법칙과 키르히호프 법칙에 의해 각 열의 뉴런에 자동으로 전기 신호가 계산되어 출력된다. 행렬 벡터 곱을 수행하기 위해 수 차례 메모리와 연산 장치 사이에서 데이터를 이동시켜야 하는 기존 폰 노이만 구조와 다르게 crossbar array 구조는 단 한 번의 연산으로 행렬 벡터 곱을 수행할 수 있다.

Crossbar array 이상으로 실제 뉴런에 더 가까운 회로를 설계하고자 하는 연구도 계속 진행되어 왔다. 실제 뉴런과 시냅스의 기능과 특성을 모방하기 위해서는 실제 뉴런과 시냅스가 어떻게 동작하는지 측정할 수 있는 시스템이 필요하다. 뉴런과 시냅스의 동작 방식을 측정하기 위해서는 집적도가 높고, 병렬 처리가 가능한 고민감도의 측정 시스템이 필요하다. (Sasaki, Minamisawa, Takahashi, Matsuki, & Ikegaya, 2009) 하지만 기존 시스템은 밀도와 민감도의 측면에서 한계를 가지고 있었다. (Spira & Hai, 2013) Patch-clamp 전국을 이용한 측정 방식은 뉴런의 action potentials (APs) propagation 외에도 postsynaptic potentials (PSPs) 등의 subthreshold event까지 측정할 수 있다는 장점이 있는 반면 집적도를 높이기 어려운 구조를 가지고 있어 지금까지도 병렬 처리할 수 있는 뉴런의 수가 10개에 불과했다. (Perin, Berger, & Markram, 2011) Nanoelectrode array를 제조하여 병렬 처리 능력을 높이려는 시도도 있었으나 이 경우에는 patch-clamp에 비해 민감도가 떨어져서 신경망 활동 측정에 활용할 수 없었다. (Abbott et al., 2017; Robinson et al., 2012)

2021년 J. Abbott *et al.*은 수 천개의 뉴런 활동을 기록할 수 있는, 고민감도 고집적도 nanoelectrode array에 대한 연구 내용을 발표한다. (Abbott et al., 2020) 연구진은 4,096개로 구성된 백금 전극을 탐침으로 사용하였고 전극으로 입력된 신호를 하부에 위치한 CMOS chip의 amplifier를 이용하여 증폭함으로써 민감도와 집적도 모두 높은 측정 시스템을 확보할 수 있었다. 연구진은 제작한 시스템을 활용해 1,700개 이상의 뉴런의 활동을 19분 동안 기록하였고 그 결과 300개의 시냅스 연결을 규명하였다.



Figure 1.1.2 기존 뉴로모픽 시스템의 회로 구조(좌)와 시냅스 구조(우)

J. Abbott *et al.*의 연구 결과를 통해 각 시냅스마다 서로 다른 propagation delay를 가진다는 것을 확인할 수 있다. 하지만 기존 뉴로모픽 시스템은 시냅스에 따라 다른 delay를 줄 수 있는 자체적인 시스템이 존재하지 않는다. Figure 1.1.2는 기존 뉴로모픽 시스템의 대표적인 회로 구조를 나타낸다. 기존 뉴로모픽 시스템에서 시냅스에 따라 다른 delay를 주기 위해서는 Figure 1.1.2의 좌측에 표시된 것과 같이 delay를 주기 위한 별도의 주변회로를 뉴런마다 추가해야 한다. 이는 회로의 집적도를 낮추고 복잡도를 높이는 문제점을 야기한다. 또한 delay 값을 수정하기 위해서는 주변회로를 매번 수정해줘야 하는데 이는 실질적으로 불가능하다.



Figure 1.1.3 OTS 기반 발화 지연 시스템 회로 구조(좌)와 시냅스 구조(우)

이러한 문제점에 착안하여 생물학적 뉴런을 보다 정확하게 모방할 수 있도록 본 연구에서는 새로운 시스템 회로 구조를 제시한다. Figure 1.1.3의 좌측은 새로운 시스템 회로 구조를 나타낸다. 어레이 외부에 주변회로를 추가하여 delay를 부여했던 Figure 1.1.3과 다르게 delay를 조절할 수 있는 소자가 어레이 내부에 위치하기 때문에 Figure 1.1.3에서는 각 시냅스별로 delay를 자체적으로 조절할 수 있고 delay 값을 수정하기에도 용이하다. 2. 발화 지연 시냅스 소자 구조 설명

2.1. RRAM 소자 동작 특성



Figure 2.1.1은 RRAM 스위칭 동작을 나타내는 모식도이다. RRAM은 metal-insulator-metal (MIM) 구조를 가지는 소자로, 절연체 양쪽에 위치한 금속은 전극 역할을 한다. 대부분의 RRAM은 metal-oxide RRAM으로(Wong et al., 2012), 소자에 전기장을 가했을 때 oxide 내에 존재하는 oxygen vacancy가 conductive filament를 형성하여 전도성을 가지게 된다.

Forming 과정 중에는 소자에 가한 높은 전기장으로 인해 soft dielectric breakdown이 일어나며 산소 이온이 insulator-metal 계면으로 이동한다. 계면으로 이동한 산소 이온은 음극 금속이 귀금속일 경우에는 non-lattice oxygen으로 환원되고 산화 가능한 금속일 경우에는 음극과 반응하여 interfacial oxide layer를 형성한다. 즉 insulator-metal 계면은 conductive filament의 구성 원소인 oxygen vacancy를 생성하는 산소 저장고로 기능한다. (Fujimoto et al., 2006)

Forming 이후 소자는 low resistance state (LRS)가 되고 이 때 소자는 conductive filament를 통해 전류가 흐른다. 소자는 RESET 과정을 통해 high resistance state (HRS)가 된다. RESET 과정 중에는 산소 이온이 음극과 절연체의 계면에서 다시 절연체로 이동하며 conductive filament의 oxygen vacancy와 재결합한다. 이 과정을 통해 Figure 2.1.1의 오른쪽 그림처럼 conductive filament의 하부에서 rupture이 일어나 소자는 HRS가 된다. 소자는 SET 과정을 통해 LRS로 돌아간다. Unipolar switching RRAM의 경우에는 RESET과 같은 polarity의 전압을 양극에 가하고 bipolar switching RRAM의 경우에는 RESET과 반대 polarity의 전압을 양극에 가한다. SET 과정에서는 다시 절연체 내에 산소 이온이 음극과 절연체의 계면으로 이동하며 conductive filament를 formation하고 소자는 LRS가 된다. RRAM 소자의 LRS는 보통 100 Ω~1kΩ, HRS는 1~10MΩ의 값을 가진다. (Chand et al., 2015; H. Lee et al., 2008; PhilipáWong, 2014) ON/OFF ratio는 10~10<sup>7</sup>의 다양한 범위를 나타내고 endurance는 보통 100 이상의 값을 보고한다. (Zahoor, Azni Zulkifli, & Khanday, 2020) 본 연구에서도 이에 상응하는 값을 목표로 RRAM 소자를 제작하고자 한다.

#### 2.2. OTS 소자 동작 특성

Resistive switching memory 의 상용화를 위해서는 적절한 selector 를 활용해야 한다. (Govoreanu et al., 2017) Selector 소자는 nA 범위의 낮은 OFF leakage 와 10MA/cm<sup>2</sup> 이상의 높은 ON current 그리고 10<sup>4</sup> 이상의 nonlinearity 를 가져야 한다. OTS 소자는 이러한 조건을 모두 만족하여 OTS 소자에 대한 연구가 활발하게 진행되어 왔다. OTS 소자는 resistive switching memory array 제작 시 selector 의 후보로서 연구가 진행되어 왔다. (Ovshinsky, 1968)

하지만 아직 OTS 소자의 switching 메커니즘에 대해서는 통일된 의견이 형성되지 않았다. 문턱 전압 이하에서 흐르는 sub-threshold 전류에 대해서는 Poole, Poole-Frankel 또는 amorphous matrix 의 tail/gap states 와 mobility gap 사이에 thermally assisted tunneling/hopping 으로 설명되어 왔다. 문턱 전압 이상의 전압에서는 이 states 들이 전자로 채워지면서 흐르는 전류가 급격하게 증가하게 된다. (Calderoni, Ferro, Ielmini, & Fantini, 2010; Ielmini, 2008; Ielmini & Zhang, 2007) OTS 소자의 전류가 흐르는 방식을 filamentary switching 으로 설명한 연구 또한 존재한다. (Chai et al., 2019; I. Karpov et al., 2007; V. Karpov, Kryukov, Karpov, & Mitra, 2008)

GeSe 계 임계스위칭 소자의 동작 원리에 대해서는 Clima et al.의 연구에 잘 요약되어 있다. (Clima et al., 2017) 스퍼터링에 의해 증착된 GeSe 물질은 비결정질 결정 상태를 가진다. 이 때 GeSe 물질은 GeSe<sub>2</sub>와 GeSe 의 혼합 상태인데 GeSe 는 GeSe<sub>2</sub>에 비해 dangling bond 를 가지기 때문에 더 높은 band gap 을 가진다. 한편 두 상태는 electric field 에 의한 가열로 인해 상호 변환 가능하다. (Epstein, 1926) 이러한 원리로 GeSe 계 임계스위칭 소자는 전기장 하에서 전도성이 바뀌게 된다.

OTS는 일반적으로 10<sup>6</sup>의 endurance, 10<sup>8</sup>의 on/off ratio, 10<sup>4</sup>의 selectivity 의 특징을 가진다. (Verdy et al., 2017)

7

#### 2.3. 발화 지연 시냅스 소자 동작 특성

Figure 1.1.3의 우측은 OTS 기반 발화 지연 시냅스 소자 구조를 나타낸다. Figure 1.1.3의 시냅스 구조와의 차이점을 빨간색 상자로 표시하였다. Figure 1.1.3의 구조에서는 어레이의 스위치에 해당하는 트랜지스터에 OTS 스위치가 직렬 연결되었다. 또한 이 OTS 스위치에 가변저항을 직렬 연결하였다. 이 가변저항은 OTS 스위치를 프로그래밍하는 데에 쓰인다. OTS 스위치는 인가한 전압과 switching delay 사이에 지수 반비례 관계가 존재한다. (S. Lee, Yoo, Park, & Hwang, 2020) OTS 스위치와 연결된 가변저항의 값을 조절함에 따라 voltage divider law에 따라 OTS 스위치에 가해지는 전압을 조절할 수 있다. 이에 따라 OTS 스위치의 switching delay가 바뀌고 결과적으로 시냅스에 주어지는 delay를 조절할 수 있다.



Figure 2.3.1 OTS 기반 발화 지연 시냅스 소자 구조(좌)와 서로 다른 가변저항 값에 따라 변화하는 delay 모식도(우)

Figure 2.3.1는 OTS 기반 발화 지연 시냅스 소자에서 delay를 조절하는 방식을 상술한다. 설명을 위해 서로 다른 crossbar array column 1, 2에 각각 가변저항 R<sub>delay1</sub>과 R<sub>delay2</sub>이 있다고 가정한다. 이 때 R<sub>delay2</sub>가 R<sub>delay1</sub>보다 크다고 가정하자. 그러면 voltage divider law에 따라 column 2의 OTS 스위치에 걸리는 전압이 column 1의 OTS 스위치에 걸리는 전압보다 작다. 한편 OTS 스위치와 switching delay 사이에는 지수 반비례 관계가 성립하기 때문에 column 2의 OTS 스위치의 switching delay가 더 크다. 결과적으로 column 2에 가해지는 delay가 더 길게 된다.

OTS 스위치에 직렬 연결하는 가변저항은 다음과 같은 조건을 만족해야 한다. 첫째 구조가 간단하고 제작이 용이해야 한다. 가변저항은 최종적으로 어레이에 포함될 소자로, 구조가 복잡하면 어레이와 통합하기 어렵기 때문이다. 둘째 높은 scalability를 가져야 한다. 마찬가지로 어레이에 통합하기 위해서는 소자의 단위 면적이 작아야 하기 때문이다. 셋째 비휘발성 소자여야 한다. 어레이가 동작하는 동안 가변저항은 기존 전기전도도를 계속 유지해야 하기 때문이다. 마지막으로 multilevel programming이 가능해야 한다. 최종적으로 여러 종류의 delay를 구현하기 위해서는 binary 소자보다는 multilevel programming이 가능한 소자가 활용성이 더 높기 때문이다. 이러한 조건을 모두 만족하는 소자는 RRAM이다. RRAM은 단순한 MIM 구조를 가지고, 4F<sup>2</sup>의 작은 단위 면적을 가지는 비휘발성 메모리이다. (Wong et al., 2012) 또한 RRAM은 multilevel programming을 발표한 연구 이미 다수 존재하기 때문에 차후에 여러 종류의 delay를 구현할 가능성도 높다. (Jang, Park, Burr, Hwang, & Jeong, 2015; Le et al., 2018; Liu, Gao, Wu, Tian, & Ren, 2021; Park et al., 2015; Prakash et al., 2014; Sheng et al., 2019)

결론적으로, 본 연구는 뉴런마다 다른 delay를 나타내는 실제 생물학적 뉴런에 더 가까운 회로를 제작하기 위해 OTS switch와 가변저항 RRAM을 직렬 연결한 소자를 제작한다. 그리고 이러한 OTS 기반 발화 지연 시냅스 소자가 정상적으로 동작하는지 확인하는 것이 본 연구의 목표이다.

## 3. 실험 및 분석 방법

#### 3.1. RRAM 소자 제작 및 측정

#### 3.1.1. RRAM 소자 제작

RRAM 소자는 MIM 구조를 가진다. Bottom electrode는 100 nm dry oxidation wafer 위에 증착한다. Dyr oxidation SiO2를 bottom electrode 밑에 배치하는 이유는 두 가지이다. 첫째 bottom electrode에서 Si substrate로 흐르는 전류를 차단하기 위해 insulating layer가 필요하기 때문이다. 둘째 SiO2를 산화 과정을 통해 형성하면 root mean square (RMS) surface roughness 0.2 nm 미만의 우수한 막질을 가진 insulating layer를 형성할 수 있기 때문이다. (Shi et al., 2015)

Bottom electrode 물질은 TiN을 사용한다. TiN은 기존 CMOS 공정과 compatible한 전도성 합금으로 RRAM 제작에 실제로 가장 많이 활용되는 물질이다. (Lanza et al., 2019) Switching layer 또한 같은 이유로 HfO2를 선택하였다. (Wilk, Wallace, & Anthony, 2001) Top electrode는 Pt를 사용하였다.



Figure 3.1.1.1 RRAM 단일 소자 구조

Figure 3.1.1.1은 본 연구에서 RRAM 특성 최적화를 위해 제작한 RRAM 단일 소자 구조이다. 본 연구에서는 Figure 3.1.1.1와 같이 bottom electrode를 소자 간에 공유하는 common bottom electrode를 이용했는데 그 이유는 이 구조가 가장 단순한 구조여서 특성 최적화를 위해 여러 물질을 screening하기에 가장 적합하기 때문이다. Common bottom electrode를 이용하여 특성 최적화를 확인한 후 OTS 소자와 통합된 소자를 만들 예정이다.



Figure 3.1.1.2은 RRAM 단일 소자의 top electrode를 패터닝하기 위해 사용한 쉐도우 마스크이다. 쉐도우 마스크의 크기는 21 mm × 21 mm이고 100 µm 직경의 원형이 0.5 mm 간격으로 패터닝 되어 있다. Figure 3.1.1.2의 쉐도우 마스크를 이용하면 총 19 × 19의 top electrode를 한 소자 내에 패터닝할 수 있다. Table 3.1.1.1은 RRAM에 사용된 물질, 증착 장비, 박막 두께를 나타낸다.

Material	Device	Thickness
Pt	Evaporator	100 nm
HfO <sub>2</sub>	ALD	7 nm
TiN	Sputter	100 nm
SiO <sub>2</sub> /Si	Furnace	100 nm

Table 3.1.1.1 RRAM에 사용된 물질, 증착 장비, 박막 두께

3.1.2. RRAM 소자 측정



Figure 3.1.2.1 RRAM DC I-V sweep 측정 방식

Figure 3.1.2.1은 RRAM 측정 방식을 나타낸다. 먼저 conductive filament 형성을 위해 compliance current (CC)를 설정한 상태로 강한 전압을 인가하여 소자를 forming 시킨다. 다음으로 CC를 해제한 상태로 forming의 반대 polarity 전압을 가해 소자를 RESET 시킨다. 마지막으로 CC를 설정한 상태로 RESET의 반대 polarity 전압을 가해 소자를 SET 시킨다.



Figure 3.1.2.2 RRAM pulse test parameter

한편 실제 chip 동작을 고려하여 pulse test 또한 진행하였다.

Pulse는 pulse generator로 인가하고 저항은 eg4155로 계산하며 NI switch와 연구실의 LabView 모듈을 이용하여 펄스 방향과 횟수를 조절하였다.

소자가 안정적으로 동작하는 최적 펄스를 찾기 위해 rise/fall time, pulse width, pulse amplitude를 바꿔가며 실험을 진행했다.

# 3.2. OTS 소자 제작 및 측정

## 3.2.1. OTS 소자 제작



Figure 3.2.1.1 N-doped GeSe계 OTS 소자 모식도(좌)와 소자 단면 SEM(우)

Figure 3.2.1.1의 좌측 그림은 N-doped GeSe계 OTS 소자 모식도를 나타낸다. OTS 소자 또한 RRAM과 마찬가지로 초기에는 쉐도우 마스크를 이용한 간단한 구조로 제작하여 임계 스위칭 층의 조성에 따른 특성 최적화를 하고자 한다. 임계 스위칭 층은 Ge-Se 이성분계 칼코제나이드 물질을 Ge<sub>0.3</sub>Se<sub>0.7</sub> 타겟과 Ge 타겟을 통해 Ge:Se 조성비를 조절하여 제작한다. 본 연구팀이 보유하고 있는 마그네트론 스피터링 시스템을 이용하여 제작하였으며, 모든 증착 층의 공정 최적화를 진행하였다.

## 3.2.2. OTS 소자 측정

OTS 소자도 RRAM 소자와 마찬가지로 DC I-V sweep을 통해 특성을 확인하고, 실제 chip의 동작 방식을 고려하여 pulse test를 진행했다. 다만 OTS 소자의 경우 전기적 특성과 delay 양상을 측정하기에 더 용이한, B1500A 반도체 파라미터 분석기의 WGFMU (Waveform Generator/Fast Measurement Unit) 모듈을 통해 측정하였음.

# 3.3. RRAM+OTS 소자 제작 및 측정

### 3.3.1. RRAM+OTS 소자 제작



Figure 3.3.1.1 RRAM+OTS 소자 마스크 도면 일부

Figure 3.3.1.1은 발화 지연 시냅스 소자를 제작하기 위해 디자인한 마스크 중 소자 부분을 확대한 결과이다. Figure 3.3.1.1 상단에 표시된 50 µm는 RRAM과 OTS의 active area의 길이를 나타낸다. 소자의 우측 상단에 RRAM, 소자의 좌측 상단에 OTS를 위치하였다. 소자의 RRAM과 OTS의 top electrode를 분리하여 각 소자의 동작 상태를 확인할 수 있도록 설계하였다. 즉, RRAM의 정상 동작을 확인하기 위해서는 bottom electrode를 접지한 상태에서 RRAM top electrode에 전기 신호를 주면 되고 OTS의 정상 동작을 확인하기 위해서는 bottom electrode를 접지한 상태에서 OTS top electrode에 전기 신호를 주면 된다. RRAM과 OTS를 직렬 연결한 측정 결과를 확인하려면 RRAM의 TE를 접지한 상태로 OTS의 top electrode에 전기 신호를 주면 된다.



Figure 3.3.1.2 RRAM+OTS 소자 마스크 도면 전체

Figure 3.3.1.2는 RRAM+OTS 소자 마스크 도면 전체를 나타낸다. Active area를 10 µm × 10 µm, 50 µm × 50 µm, 100 µm × 100 µm로 split하였고 각 active area별로 네 개의 동일한 소자를 제작했다. 마스크의 양쪽에는 align key를 위치하였다. 마스크 가장 하단에는 각 박막 층의 두께를 확인할 수 있도록 line and space 패턴을 위치하였다. 우측 하단에는 각 전극의 면저항을 확인할 수 있도록 TLM 패턴을 위치하였다. 이외에 소자의 동작 확인을 위해 단일 소자 패턴을 위치하였다.



Figure 3.3.1.3는 RRAM+OTS 소자 process flow를 나타낸다.

## 3.3.2. RRAM+OTS 소자 측정

RRAM+OTS 소자도 RRAM, OTS 소자와 마찬가지로 DC I-V sweep과 pulse test를 진행할 예정이다.

## 4. 결과 및 논의

#### 4.1. RRAM 소자 제작 및 측정

## 4.1.1. RRAM 단일 소자 switching 성능 개선 실험

실험 초기에 RRAM 단일 소자는 낮은 endurance와 yield를 보였다. 최대 endurance는 94 cycles에 불과했고 소자의 동작 확률도 10%에 머물렀다. 실험 초기 RRAM 소자의 구성은 Pt 100 nm / HfO<sub>2</sub> 10 nm / TiN 100 nm였다.

소자의 성능을 개선하기 위해 switching 불량 원인을 분석했다. 이를 위해 먼저 RRAM의 switching 메커니즘을 분석했다. (N. Xu et al., 2008) RRAM의 LRS와 HRS에서 전류는 conductive filament의 oxygen vacancy 사이에서 electron hopping을 통해 흐른다. LRS와 HRS 사이에 switching은 conductive filament의 formation과 일어난다. Conductive filament의 formation과 rupture를 통해 rupture에 주된 영향을 미치는 변인은 electric field이다. Electric field의 크기는 소자에 가해진 전압을 소자의 두께로 나누어서 구할 수 있다. RESET은 conductive filament의 말단에 있는 oxygen vacancy가 non-lattice oxygen ion과 재결합하여 일어나고 SET은 oxygen ion이 전기장에 의해 lattice에서 분리되며 oxygen vacancy를 형성하는 dielectric soft breakdown이다. Switching failure는 주로 non-lattice oxygen ion이 충분히 존재하지 않아서 일어난다. 결론적으로 switching을 개선하기 위해서는 switching layer의 두께와 전압의 amplitude를 조절하여 전기장의 세기를 조절하거나 switching layer의 두께를 조절하여 non-lattice oxygen ion의 양을 바꾸는 방법을 사용할 수 있다.

21





Figure 4.1.1.1 RRAM 단일 소자 switching 불량 원인 분석

Figure 4.1.1.1은 실제로 소자에서 일어나는 switching 불량 형태를 분류한 결과이다. 분류 결과 크게 불량은 네 가지 형태로 일어남을 확인할 수 있었다. 첫 번째 형태는 forming을 하기 전에 이미 소자가 LRS 상태이고 HRS로 돌아오지 않는 경우이다. 이를 해결하기 위해 소자 구조를 개선하는 방법을 사용할 수 있다. 현재 소자 구조는 bottom electrode 외에도 switching layer를 공유하는 구조이기 때문에 인접한 소자의 switching이 선택한 소자의 switching에 영향을 줄 수 있다. 즉 다른 소자를 switching할 때 선택하지 않은 소자의 switching을 야기하여 breakdown을 일으킬 수 있다. 소자 구조를 개선하는 방법 외에 전기장의 크기를 줄이는 방법을 택할 수 있다. 이를 위해 switching layer의 두께를 키우거나 동작 전압과 CC를 낮출 수 있다. 두 번째 형태는 forming 이후에 breakdown이 일어나는 경우이다. 이 경우도 첫 번째 경우와 거의 마찬가지의 해결책을 갖는다. Switching layer의 두께를 키우거나 동작 전압과 CC를 낮춰 전기장의 크기를 줄인다면 문제를 해결할 수 있을 것이다. 세 번째 경우는 forming이 아예 안 일어나는 경우이다. 이 경우에는 소자에 가해지는 전기장의 크기가 너무 작아서 switching이 일어나지 않는다고 생각할 수 있다. 그러므로 switching layer 두께를 줄여서 문제를 해결할 수 있다. 동작 전압이나 CC를 키우는 방법은 최종적으로 소비 전력이나 소자의 안정성을 낮추므로 지양해야 한다. Forming이 일어나지 않는 방법을 해결하기 위해 단순히 switching layer의 두께를 줄이는 방법 외에도

contact resistance를 낮추는 방법이 있다. 소자의 두께나 전압의 크기보다 contact resistance가 너무 높아서 소자에 전달되는 전압이 낮을 수 있기 때문이다. Contact resistance를 낮추기 위해서는 bottom electrode 쪽 TiN/SiO<sub>2</sub> 또는 top electrode 쪽 Pt/HfO<sub>2</sub> 계면을 개선하는 방법을 생각할 수 있다. TiN/SiO2 계면을 개선하기 위해 공정을 시작하기 전에 조각 시편을 acetone, IPA, DI water 내에서 sonicator로 클리닝하는 과정을 추가하였다. 또한 Pt/HfO2 계면을 향상하는 방법으로는 Pt를 증착하기 전에 adhesion layer로 Ti를 추가하는 방법이 있다. Pt는 oxide에 바로 증착하면 계면이 안 좋다는 연구 결과가 있기 때문이다. (Abe et al., 2003) 마지막 경우는 RESET 도중 breakdown이 일어나는 경우이다. 보통 RESET 중에는 oxygen ion이 oxygen vacancy와 재결합해야 하지만 switching 중 발생하는 Joule heating으로 인해 hot carrier가 형성되어 더 굵은 filament가 형성되며 breakdown이 일어나는 것으로 보인다. 이러한 현상이 일어나는 주된 이유는 switching layer 내에 non-lattice oxygen ion이 부족하여 bipolar RRAM 대신 unipolar RRAM이 형성되었기 때문이다. 이를 개선하기 위해서 bottom electrode 쪽 TiN/HfO<sub>2</sub> 계면에 defect가 더 많이 형성될 수 있도록 TiN 증착 전 O<sub>3</sub> plasma를 가하는 방식을 생각할 수 있다. 다른 방법으로는 Joule heating이 과도하게 커지지 않도록 동작 전압, CC를 감소하거나 switching layer의 두께를 키우는 방법을 생각할 수 있다.

Switching 불량 형태를 네 가지 경우로 분류했을 때 공통적인 해결책은 동작 전압이나 CC를 조절하는 방법 또는 switching layer의 두께를 조절하는 방법이 있었다. 따라서 switching layer의 두께를 키우거나 줄이며 CC를 조절하여 switching 형태가 어떻게 바뀌는지 관찰하는 방식으로 실험을 진행하였다.





Figure 4.1.1.2 HfO<sub>2</sub> 두께 증가 후 측정 결과 (10 nm → 15 nm)

Figure 4.1.1.2는 HfO<sub>2</sub> switching layer의 두께를 10 nm에서 15 nm로 키우며 CC를 100 µA에서 1 µA로 조절하며 I-V curve를 얻은 결과이다. 결론적으로 성능은 저하되었다. 성능 저하는 크게 세 가지 형태로 관찰되었다. 첫째 forming voltage가 증가하였다. 이는 HfO<sub>2</sub>의 두께가 증가하며 소자에 가해지는 전기장의 크기가 감소했기 때문에 당연한 결과이다. 특기할 만한 점은 forming voltage의 증가 폭이 HfO<sub>2</sub> 두께의 증가 폭과 같다는 점이다. 이는 switching layer의 두께 차이로 인한 전기장의 크기 변화가 switching 불량의 주된 원인이 맞다는 사실을 반증한다. 둘째 RESET 중 breakdown이 일어났다. 이는 Figure 4.1.1.2의 세번째 경우에서 볼 수 있듯이 RESET voltage가 증가하였기 때문이다. RESET voltage가 증가하였으므로 RESET 중 breakdown이 일어날 확률 또한 높아졌다고 생각할 수 있다.



Figure 4.1.1.3 HfO₂ 두께 감소 후 endurance 측정 결과 (10 nm → 7 nm)

Figure 4.1.1.3는 HfO<sub>2</sub> switching layer의 두께를 10 nm에서 7 nm로 줄였을 때 측정 결과이다. 결과적으로 성능은 개선되었다. Endurance 값은 기존 94 cycles에서 1,000 cycles 이상으로 크게 개선되었다. ON/OFF ratio 또한 10<sup>3</sup>의 높은 수치를 보였고 LRS와 HRS는 각각 2 kΩ과 2 MΩ이었다.



Figure 4.1.1.4 HIO2 十川 石立 十 yield test (10 lilli ラ / lilli)

Figure 4.1.1.4는 HfO<sub>2</sub> 두께를 10 nm에서 7 nm로 줄인 후 yield test를 진행한 결과이다. 임의의 소자 10개를 선택하여 10 cycles를 확인한 결과 전부 정상적으로 동작하여 동작 확률이 기존 10%에서 100%로 크게 개선되었음을 확인할 수 있었다.

### 4.1.2. RRAM 단일 소자 pulse test

Pulse width ( <sup>ns</sup> )	Rise/fall time ( <sup>ns</sup> )	Pulse amplitude (V)
100	10	6.5

Table 4.1.2.1 RRAM pulse test 최적 펄스 조건

RRAM pulse test를 진행하기 위해 pulse width, rise/fall time, pulse amplitude에 대해 RRAM switching이 가장 잘 일어나는 최적 펄스 조건을 찾았다. Table 4.1.2.1은 그 결과 얻은 최적 펄스 조건이다. Pulse width 100 ns, rise/fall time 10 ns, pulse amplitude 6.5 V이 최적 조건임을 확인할 수 있다.



Figure 4.1.2.1 RRAM switching cycle에 따른 SET, RESET 저항 값

Figure 4.1.2.1은 최적 펄스 조건으로 RRAM switching cycle을 테스트 한 결과 SET, RESET 저항 값을 나타낸다. 소자의 ON/OFF ratio는 10 이상으로 확인되었고 endurance는 84 cycles로 확인되었다.

DC I-V sweep과 비교했을 때 오히려 endurance가 열화된 결과를 확인할 수 있다. 이는 크게 두 가지 이유로 생각되는데 첫번째 이유는 pulse test 측정 시스템에서는 CC가 존재하지 않기 때문이다. CC가 존재하지 않으면 소자에 과도한 전류가 흐르기 때문에 더 빠르게 소자가 열화될 가능성이 있다. 이를 해결하기 위해 소자에 직렬 저항 패턴을 추가하는 방법을 도입할 수 있다. 두번째 이유는 소자의 구조이다. 이 소자 또한 bottom electrode를 공유하기 때문에 소자의 특성이 나쁠 수 있다. 이를 해결하기 위해 포토 공정을 도입하여 bottom electrode를 분리하는 방법을 도입할 수 있다.

## 4.2. OTS 소자 제작 및 측정

## 4.2.1. N-doped GeSe OTS 소자 제작 및 측 정

GeSe계 임계스위칭 소자는 첫 스위칭 (First Fire, FF)에는 높은 임계전압에 의해 발생하며, 그 이후 스위칭에서는 비교적 낮은 임계전압에서 스위칭이 일어나는 것으로 알려져 있다. 본 특성은 triangular pulse를 인가하여 확인할 수 있었다.



Figure 4.2.1.1 Triangular pulse FF 및 후속 스위칭

Figure 4.2.1.1는 N-doped GeSe OTS 소자에 triangular pulse를 인가하였을 때의 첫 스위칭과 후속 스위칭의 측정 결과이다. 상술한대로 첫 스위칭 이후로는 비교적 낮은 임계전압에서 스위칭이 일어나는 것을 확인할 수 있다. 임계전압은 평균적으로 약 1.8 V, 홀딩전압은 약 1.5 V, ON/OFF ratio는 10<sup>3</sup>-10<sup>4</sup>의 특성을 확보할 수 있었음.

임계 스위칭 소자의 메커니즘은 E-field induced filamentary

형성이라는 스위칭 메커니즘으로 최근 학계에서 발표된 바가 있다. (Chai et al., 2019) 인가된 E-field에 따른 스위칭 delay는 식(1)과 같고, 식으로부터 알 수 있듯이 switching delay는 인가전압과 지수 반비례 관계를 갖는다. (S. Lee et al., 2020)



$$\tau_d = \tau_0 \exp\left(\frac{W(E)}{kT}\right) = \tau_0 \exp\left(\frac{W_0 \alpha^{\frac{3}{2}} E_0 d}{kT}\right) \tag{1}$$

Figure 4.2.1.2 Square pulse 인가전압 변화에 따른 delay 변화 측정 결과

임계 스위칭 소자의 FF 및 후속 임계 스위칭을 확인한 후에는 square pulse를 인가하여 switching delay를 측정하는 실험을 진행하였다. Figure 4.2.1.2는 그 결과를 나타낸다. 1.7 V에서는 3.06 μs, 2.0 V에서는 1.90 μs의 delay를 보여 전압과 delay 사이에 반비례 관계를 확인할 수 있다.



Figure 4.2.1.3 Square pulse 인가전압 변화에 따른 delay 변화 측정 결과

식 (1)의 지수 반비례 관계를 확인하기 위해 Square pulse의 전압 크기를 2.0 V에서 3.0V 범위에서 바꿔 가며 delay를 반복 측정하였다. Figure 4.2.1.3은 그 결과를 나타낸다. 인가전압에 따른 평균 switching delay는 1 μs에서 10 μs 사이에 값을 가지는 것을 확인할 수 있다. 또한 측정된 switching delay 값들은 식 (1)의 인가전압 간의 관계가 잘 일치하는 것을 확인할 수 있었다.

#### 4.2.2. N-doped AgGeSe OTS 소자 제작 및 측정

N-doping과 pulsed test를 통해 OTS 소자가 정상 동작함을 확인할 수 있었고 endurance, 임계전압, on-off ratio 등의 특성을 측정할 수 있었다. 그러나 본 Ge<sub>0.3</sub>Se<sub>0.7</sub> 조성의 임계 스위칭 소자에서 확인된 switching delay는 1-10 µs로 ms 범위에 있는 실제 뉴런 발화 지연시간 구현에 비해 너무 짧은 작동 범위를 가지고 있다. 따라서 다른 조성에서의 임계 스위칭 소자를 제작하여 switching delay를 개선을 시도하였다.

연구를 진행하기에 앞서 먼저 GeSe계 임계 스위칭 소자의 switching mechanism 분석을 통해 GeSe계 스위칭 소자의 switching 시간이 짧은 이유를 파악했다. (Clima et al., 2017) GeSe계 임계 스위칭 소자는 임계 스위칭 시, Ge의 non-bonding orbital인 valencealternating-pair(VAP) 간의 interaction으로 인해 발생하는 전기장에 의해 달라진 bonding configuration(C<sup>+</sup><sub>3</sub>, C<sup>-</sup><sub>1</sub>)이 local한 영역에서 trapping state를 형성하며 순간적인 전기전도도의 변화를 가져오는 것으로 알려져 있다. 즉 전자 orbital의 변화로 인해 local electrical path가 형성되는데 이는 electronic response이기 때문에 매우 짧은 응답 속도를 보일 것으로 예상할 수 있다. 이러한 메커니즘으로 인해 GeSe계 임계 스위칭 소자의 delay가 짧다고 설명할 수 있다.

이러한 GeSe계의 electronic response로 인한 짧은 delay를 더 길게 할 수 있도록 atomic diffusion을 고려한 조성을 시도했다. Ag<sup>+</sup> ion은 GeSe를 solid electrolyte로 활용하여 GeSe 내에서 ionic conductivity가 좋다고 알려져 있다. (Kawasaki, Kawamura, Nakamura, & Aniya, 1999) 이러한 점에 착안하여 N-doped GeSe에 Ag를 cosputtering 하여 AgGeSe 소자의 제작을 시도하였다.

34



Figure 4.2.2.1 N-doped AgGeSe 소자 동작 모식도

Figure 4.2.2.1 는 N-doped AgGeSe의 동작 모식도를 나타낸다. 먼저 top electrode에 positive bias를 인가하면 switching layer에 cosputter 된 Ag가 산화되어 diffusive Ag+ ion이 생성된다. 생성된 Ag<sup>+</sup> ion은 인가된 전기장으로 인해 bottom electrode로 diffusion한다. Bottom electrode 주변에 모인 Ag<sup>+</sup> ion은 환원되어 filament 구조를 형성한다. 이는 RRAM의 forming과 유사한 동작으로, conductive 한 영역을 만들어 GeSe계의 임계 스위칭을 assist할 수 있는 소자 구조이다. 또한 GeSe layer에 heavy N-doping을 통해 leakage를 낮추고 임계 전압을 높이는 효과를 주어 Ag<sup>+</sup> 이온의 diffusion 없이는 동작할 수 없도록 설계하였다.



Figure 4.2.2.2 N-doped AgGeSe SEM 단면과 EDS 성분 분석

Figure 4.2.2.는 N-doped AgGeSe 소자의 SEM 단면 결과와 EDS 성분 분석이다. SEM 단면 결과로부터 의도한 두께대로 소자가 잘 제작된 것을 확인할 수 있다. 또한 Figure 4.2.2.2의 우측에는 EDS 성분 분석 결과를 표시했다. AgGeSe switching layer는 Ag 타겟과 Ge<sub>0.3</sub>Se<sub>0.7</sub> 타겟을 이용해서 증착했는데 분석 결과에서 co-sputtering이 잘 진행되었음을 확인할 수 있다. 또한 Ag<sub>2</sub>Se가 열역학적으로 매우 안정한 phase임을 고려했을 때 (Fischer-Colbrie, Bienenstock, Fuoss, & Marcus, 1988) 증착된 AgGeSe는 Ag<sub>2</sub>Se와 GeSe가 1:1 stoichiometry로 증착된 것을 역산할 수 있다. Ag<sub>2</sub>Se가 열역학적으로 안정하므로 외부 전기장으로 인해 형성된 Ag<sup>+</sup> 이온은 다시금 주변 Se atom에 solute하여 Ag<sub>2</sub>Se 상태가 될 것으로 예측된다. 이러한 전기장 유무에 따른 ion 형성 및 안정상으로의 solvation으로 인해, 기존 GeSe계 임계 스위칭 소자의 electronic 메커니즘과 달리 atomic diffusion이 동반될 수 있어 기존보다 느리고 폭 넓은 스위칭 지연시간을 기대해볼 수 있다.



Figure 4.2.2.3 AgGeSe 임계 스위칭 소자와 GeSe 스위칭 소자의 delay 비교

Figure 4.2.2.3는 AgGeSe 소자와 GeSe 소자의 delay를 비교한 결과이다. 예측한대로 AgGeSe 소자의 delay가 GeSe 소자의 delay보다 긴 것을 확인할 수 있다.



Figure 4.2.2.4 AgGeSe 임계 스위칭 소자 최대 delay

Figure 4.2.2.4는 AgGeSe 임계 스위칭 소자에서 관찰된 최대 delay다. 최대 delay는 4.8 ms로 기존 GeSe계 임계 스위칭 소자와 비교했을 때 100배 이상 증가한 수치이고 실제 뉴런의 발화 지연 시간과도 비슷한 수준의 수치이다. 이로부터 AgGeSe 소자를 이용하면 실제 뉴런의 발화 지연과 유사한 특성을 얻을 수 있다고 유추할 수 있다.



Figure 4.2.2.5 AgGeSe 임계 스위칭 소자 DC I-V sweep

Figure 4.2.2.5는 AgGeSe 임계 스위칭 소자의 DC I-V sweep 측정 결과이다. 소자의 임계 전압은 약 1.6 V, 홀딩 전압은 약 1.2 V, ON/OFF ratio는 10<sup>4</sup>로 확인되었다.

#### 4.3. RRAM+OTS 소자 제작 및 측정

#### 4.3.1. RRAM+OTS 소자 제작



Figure 4.3.1.1 RRAM+OTS 소자 광학 현미경 촬영 이미지

오류! 참조 원본을 찾을 수 없습니다.는 RRAM+OTS 소자 중 active area 50 µm × 50 µm에 해당하는 부분을 광학 현미경으로 촬영한 결과이다. 좌측 pad가 OTS의 top electrode에 해당하고 우측 pad가 RRAM의 top electrode에 해당한다.

OTS의 active area 부분이 우측의 RRAM과 다르게 깔끔하지 못하다는 것을 **오류! 참조 원본을 찾을 수 없습니다.**에서 확인할 수 있다. 공정 장비 상의 문제로 GeSe layer를 증착한 이후에 수 일간 실험을 진행하지 못했는데 이 과정에서 GeSe layer의 열화가 일어났을 수 있다고 생각한다. 또한 etch gas로 SF<sub>6</sub>를 이용했는데 SF<sub>6</sub> gas의 etching damage가 컸기 때문일 수도 있다. 보다 개선된 소자 성능을 위해서는 GeSe layer etch 또한 최적화를 더 진행해야 할 것이다.

한편 bottom electrode 또한 색깔이 기존의 TiN 색깔과 다른 것을 확인할 수 있다. TiN의 기존 색깔은 OTS의 top electrode 색깔로부터 알 수 있다. Bottom electrode는 공정의 가장 처음에 증착한 이후 다른 layer의 공정을 진행할 때 항상 노출되어 있다. 이 과정에서 식각이 계속 진행되어 TiN이 남아있지 않거나 etch로 인한 데미지를 받았을 수 있다고 생각된다. SEM 촬영과 I-V sweep을 통해 이 가설을 확인하고 다른 layer의 etch rate을 보다 정확하게 밝힐 필요가 있다.

#### 4.3.2. RRAM+OTS 소자 측정

측정 결과 소자에 전류가 흐르지 않는 것을 확인하였다. RRAM+OTS 소자 외에도 시험 용도로 만든 단일 소자 역시 동작하지 않았다.

예상되는 원인은 bottom electrode가 다른 layer의 포토 공정 중에 노출되어 TiN이 전부 식각 되거나 etch damage로 인해 막질이 악화되었을 수 있다. 이러한 문제를 해결하기 위해서는 우선 SEM 촬영과 I-V sweep을 통해 pad의 전도성을 확인하고 이후 각 layer의 etch rate를 보다 정확하게 밝혀야 한다.

## 4. 결론

본 연구에서는 뉴런 별로 다른 delay를 부여할 수 없는 기존 시냅스 회로의 한계점에 착안하여 뉴런 별로 다른 delay를 부여할 수 있는 새로운 시냅스 회로 구조를 제시하였다. OTS switching delay를 활용한 RRAM 기반 발화 지연 시냅스 소자를 활용하면 뉴런마다 다른 delay를 부여할 수 있다.

발화 지연 시냅스 소자를 제작하기에 앞서 RRAM과 OTS 단일 소자의 특성 최적화를 진행했다. 특성 최적화를 위해 두 소자 모두 간단한 공정으로 제작이 가능한 common bottom electrode 구조를 사용하였다.

RRAM은 동작 확률 10%, endurance는 94 cycles로 불량한 switching 특성을 보였다. 이를 개선하기 위하여 RRAM의 switching 메커니즘을 분석한 결과 switching layer의 두께, forming voltage, compliance current를 조절하여 switching 특성을 개선할 수 있음을 파악했다. Switching layer 두께를 10 nm에서 7 nm로 줄인 결과 동작 확률을 100%, endurance cycle는 1000 cycles 이상으로 크게 개선할 수 있었다.

OTS 소자는 GeSe계 switching layer를 이용한 결과 1-10 µs의 delay를 보였다. 한편 실제 뉴런의 발화 지연 시간은 수 ms에 달하기 때문에 delay를 늘릴 필요가 있었다. Delay를 늘리기 위해 GeSe계 소자의 switching 메커니즘을 분석하였다. 분석 결과 GeSe계 소자에서는 electron orbital의 interaction으로 인해 switching이 일어나고 이는 electronic response로 반응 시간이 짧기 때문에 delay 또한 짧다는 사실을 알 수 있었다. 이를 개선하기 위해 Ag를 cosputtering하여 atomic diffusion process를 추가하여 delay를 늘리는 방법을 시도하였다. 결과적으로 delay를 4.8 ms까지 늘릴 수 있었고 이는 실제 뉴런의 발화 지연에 상당하는 수치이다.

마지막으로 RRAM과 OTS를 통합한 소자를 제작 완료했지만 소자의 정상적인 동작은 확인할 수 없었다. 전류가 흐르지 않았는데 bottom electrode가 다른 layer를 식각하는 과정에서 손상을 받았다고 생각된다. 정확한 소자 동작을 위해서는 etch rate 등의 변수를 보다 최적화할 필요가 있다고 생각된다.

- Abbott, J., Ye, T., Krenek, K., Gertner, R. S., Ban, S., Kim, Y., . . . Ham, D. (2020). A nanoelectrode array for obtaining intracellular recordings from thousands of connected neurons. *Nature biomedical engineering*, 4(2), 232-241.
- Abbott, J., Ye, T., Qin, L., Jorgolli, M., Gertner, R. S., Ham, D., & Park, H. (2017). CMOS nanoelectrode array for all-electrical intracellular electrophysiological imaging. *Nature nanotechnology*, 12(5), 460-466.
- Abe, N., Otani, Y., Miyake, M., Kurita, M., Takeda, H., Okamura, S., & Shiosaki, T. (2003). Influence of a TiO2 adhesion layer on the structure and the orientation of a Pt layer in Pt/TiO2/SiO2/Si structures. *Japanese journal of applied physics*, 42(5R), 2791.
- Byun, K., Choi, I., Kwon, S., Kim, Y., Kang, D., Cho, Y. W., . . . Kim, S. (2022). Recent Advances in Synaptic Nonvolatile Memory Devices and Compensating Architectural and Algorithmic Methods Toward Fully Integrated Neuromorphic Chips. *Advanced Materials Technologies*, 2200884.
- Calderoni, A., Ferro, M., Ielmini, D., & Fantini, P. (2010). A unified hopping model for subthreshold current of phase-change memories in amorphous state. *IEEE Electron Device Letters*, 31(9), 1023-1025.
- Chai, Z., Zhang, W., Degraeve, R., Clima, S., Hatem, F., Zhang, J., . . . Garbin, D. (2019). Evidence of filamentary switching and relaxation mechanisms in Ge x Se 1-x OTS selectors. Paper presented at the 2019 Symposium on VLSI Technology.
- Chand, U., Huang, C.-Y., Jieng, J.-H., Jang, W.-Y., Lin, C.-H., & Tseng, T.-Y. (2015). Suppression of endurance degradation by utilizing oxygen plasma treatment in HfO2 resistive switching memory. *Applied Physics Letters*, 106(15), 153502.
- Clima, S., Govoreanu, B., Opsomer, K., Velea, A., Avasarala, N. S., Devulder, W., . . . Kundu, S. (2017). Atomistic investigation of the electronic structure, thermal properties and conduction defects in Ge-rich Ge x Se 1- x materials for selector applications. Paper presented at the 2017 IEEE International Electron Devices Meeting (IEDM).
- Epstein, P. S. (1926). The Stark effect from the point of view of Schroedinger's quantum theory. *Physical Review, 28*(4), 695.
- Fischer-Colbrie, A., Bienenstock, A., Fuoss, P., & Marcus, M. A. (1988). Structure and bonding in photodiffused amorphous Ag-GeSe 2 thin films. *Physical Review B, 38*(17), 12388.
- Fujimoto, M., Koyama, H., Konagai, M., Hosoi, Y., Ishihara, K., Ohnishi, S., & Awaya, N. (2006). Ti O 2 anatase nanolayer on TiN thin film exhibiting high-speed bipolar resistive switching. *Applied Physics Letters*, 89(22), 223509.
- Govoreanu, B., Donadio, G. L., Opsomer, K., Devulder, W., Afanas' ev, V., Witters, T., . . . Kundu, S. (2017). *Thermally stable integrated Se-*

based OTS selectors with> 20 MA/cm 2 current drive,> 3.10 3 halfbias nonlinearity, tunable threshold voltage and excellent endurance. Paper presented at the 2017 Symposium on VLSI Technology.

- He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep residual learning for image recognition.* Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
- Ielmini, D. (2008). Threshold switching mechanism by high-field energy gain in the hopping transport of chalcogenide glasses. *Physical Review B*, 78(3), 035308.
- Ielmini, D., & Zhang, Y. (2007). Analytical model for subthreshold conduction and threshold switching in chalcogenide-based memory devices. *Journal of applied physics*, 102(5), 054517.
- Indiveri, G., & Liu, S. C. (2015). Memory and Information Processing in Neuromorphic Systems. *Proceedings of the IEEE*, 103(8), 1379–1397. doi:10.1109/jproc.2015.2444094
- Jang, J.-W., Park, S., Burr, G. W., Hwang, H., & Jeong, Y.-H. (2015). Optimization of conductance change in Pr 1-x Ca x MnO 3-based synaptic devices for neuromorphic systems. *IEEE Electron Device Letters*, 36(5), 457-459.
- Karpov, I., Mitra, M., Kau, D., Spadini, G., Kryukov, Y., & Karpov, V. (2007). Fundamental drift of parameters in chalcogenide phase change memory. *Journal of applied physics*, *102*(12), 124503.
- Karpov, V., Kryukov, Y., Karpov, I., & Mitra, M. (2008). Field-induced nucleation in phase change memory. *Physical Review B*, 78(5), 052201.
- Kawasaki, M., Kawamura, J., Nakamura, Y., & Aniya, M. (1999). Ionic conductivity of Agx (GeSe3) 1− x (0≤ x≤ 0.571) glasses. Solid State Ionics, 123(1-4), 259-269.
- Lanza, M., Wong, H. S. P., Pop, E., Ielmini, D., Strukov, D., Regan, B. C., . . . Goux, L. (2019). Recommended methods to study resistive switching devices. *Advanced Electronic Materials*, 5(1), 1800143.
- Le, B. Q., Grossi, A., Vianello, E., Wu, T., Lama, G., Beigne, E., . . . Mitra, S. (2018). Resistive RAM with multiple bits per cell: Array-level demonstration of 3 bits per cell. *IEEE Transactions on Electron Devices, 66*(1), 641-646.
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- Lee, H., Chen, P., Wu, T., Chen, Y., Wang, C., Tzeng, P., ... Tsai, M.-J. (2008). Low power and high speed bipolar switching with a thin reactive Ti buffer layer in robust HfO2 based RRAM. Paper presented at the 2008 IEEE International Electron Devices Meeting.
- Lee, S., Yoo, J., Park, J., & Hwang, H. (2020). Understanding of the abrupt resistive transition in different types of threshold switching devices from materials perspective. *IEEE Transactions on Electron Devices*, 67(7), 2878-2883.
- Liu, Y., Gao, J., Wu, F., Tian, H., & Ren, T.-L. (2021). The Origin of CBRAM

With High Linearity, On/Off Ratio, and State Number for Neuromorphic Computing. *IEEE Transactions on Electron Devices, 68*(5), 2568–2571. doi:10.1109/ted.2021.3065013

- Ovshinsky, S. R. (1968). Reversible electrical switching phenomena in disordered structures. *Physical review letters*, *21*(20), 1450.
- Park, S., Chu, M., Kim, J., Noh, J., Jeon, M., Hun Lee, B., . . . Lee, B.-g. (2015). Electronic system with memristive synapses for pattern recognition. *Scientific reports*, 5(1), 1-9.
- Perin, R., Berger, T. K., & Markram, H. (2011). A synaptic organizing principle for cortical neuronal groups. *Proceedings of the National Academy of Sciences*, 108(13), 5419–5424.
- PhilipáWong, H.-S. (2014). Multi-level control of conductive nano-filament evolution in HfO 2 ReRAM by pulse-train operations. *Nanoscale, 6*(11), 5698-5702.
- Prakash, A., Park, J., Song, J., Woo, J., Cha, E.-J., & Hwang, H. (2014). Demonstration of low power 3-bit multilevel cell characteristics in a TaO x-based RRAM by stack engineering. *Ieee Electron Device Letters, 36*(1), 32-34.
- Robinson, J. T., Jorgolli, M., Shalek, A. K., Yoon, M.-H., Gertner, R. S., & Park, H. (2012). Vertical nanowire electrode arrays as a scalable platform for intracellular interfacing to neuronal circuits. *Nature nanotechnology*, 7(3), 180-184.
- Sasaki, T., Minamisawa, G., Takahashi, N., Matsuki, N., & Ikegaya, Y. (2009). Reverse optical trawling for synaptic connections in situ. *Journal of Neurophysiology*, 102(1), 636-643.
- Sheng, X., Graves, C. E., Kumar, S., Li, X., Buchanan, B., Zheng, L., . . . Strachan, J. P. (2019). Low-Conductance and Multilevel CMOS-Integrated Nanoscale Oxide Memristors. *Advanced Electronic Materials*, 5(9), 1800876.
- Shi, Y., Ji, Y., Sun, H., Hui, F., Hu, J., Wu, Y., . . . Duan, H. (2015). Nanoscale characterization of PM2. 5 airborne pollutants reveals high adhesiveness and aggregation capability of soot particles. *Scientific reports*, 5(1), 1-10.
- Shoeybi, M., Patwary, M., Puri, R., LeGresley, P., Casper, J., & Catanzaro, B. (2019). Megatron-lm: Training multi-billion parameter language models using model parallelism. arXiv preprint arXiv:1909.08053.
- Spira, M. E., & Hai, A. (2013). Multi-electrode array technologies for neuroscience and cardiology. *Nature nanotechnology*, 8(2), 83-94.
- Verdy, A., Navarro, G., Sousa, V., Noe, P., Bernard, M., Fillot, F., . . . Perniola, L. (2017). *Improved electrical performance thanks to Sb and N doping in Se-rich GeSe-based OTS selector devices.* Paper presented at the 2017 IEEE International Memory Workshop (IMW).
- Wilk, G. D., Wallace, R. M., & Anthony, J. (2001). High-κ gate dielectrics: Current status and materials properties considerations. *Journal of applied physics*, 89(10), 5243-5275.
- Wong, H. S. P., Lee, H.-Y., Yu, S., Chen, Y.-S., Wu, Y., Chen, P.-S., . . . Tsai,

M.-J. (2012). Metal-Oxide RRAM. *Proceedings of the IEEE, 100*(6), 1951-1970. doi:10.1109/jproc.2012.2190369

- Xu, N., Gao, B., Liu, L., Sun, B., Liu, X., Han, R., . . . Yu, B. (2008). *A unified physical model of switching behavior in oxide-based RRAM.* Paper presented at the 2008 Symposium on VLSI Technology.
- Xu, X., Ding, Y., Hu, S. X., Niemier, M., Cong, J., Hu, Y., & Shi, Y. (2018). Scaling for edge inference of deep neural networks. *Nature Electronics*, 1(4), 216-222.
- Zahoor, F., Azni Zulkifli, T. Z., & Khanday, F. A. (2020). Resistive random access memory (RRAM): an overview of materials, switching mechanism, performance, multilevel cell (MLC) storage, modeling, and applications. *Nanoscale research letters, 15*(1), 1-26.

## Abstract RRAM-based delay synapse using OTS switching delay

Younghoon Kim Materials Science and Engineering The Graduate School Seoul National University

With the development of artificial intelligence, research is focused on the development of circuits that mimic real biological neural networks for more efficient calculations. A crossbar array structure modeled after actual neurons and synapses is widely used, but the structure has a limitation in that it cannot give different delays for each neuron, unlike actual neurons. In this study, focusing on these limitations, we proposed a new synaptic circuit structure that can provide different delays for each neuron. In the new synaptic circuit structure, a different delay can be given to each neuron by connecting the OTS and RRAM in series and changing the variable resistance of the RRAM to change the voltage applied to the OTS. In order to create a device that integrates OTS and RRAM, characteristics of a single RRAM and OTS device were optimized. As a result of adjusting the switching layer thickness and compliance current, the performance of RRAM was significantly improved from the existing yield of 10% and endurance of 94 cycles to 100% of yield and endurance of 1000 cycles or more. By changing the existing GeSe-based switching layer to AgGeSe, the OTS device was able to increase the delay from the existing  $1-10 \ \mu s$  to 4.8 ms. This is a

numerical value corresponding to the firing delay of actual neurons. A device that integrates RRAM and OTS has been manufactured. However, the device turned out to be not working. To solve this matter fabrication process parameters such as etch rate should be identified.

Keywords : Ovonic threshold switching, Resistive random access memory, resistive switching, threshold switching, switching delay, neuromorphic synapse Student Number : 2021-22969