Leveraging Atmospheric-Pressure Spatial Atomic Layer Deposition and Machine Learning for Nanomaterial and Device Design

by

Olivia Marchione

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Mechanical and Mechatronics Engineering - Nanotechnology

Waterloo, Ontario, Canada, 2023

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Dr Yu Zou
Supervisor	Dr. Kevin Musselman
Internal Member	Dr. Yimin Wu
Internal Member	Dr. Mihaela Vlasea
Other Member(s)	Dr. Yaoliang Yu

Authors Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Olivia Marchione was the sole author for Chapters 1, 2, 3, 5, and 7 which were written under the supervision of Dr. Kevin Musselman and were not written for publication.

This thesis consists in part of four manuscripts written for publication, those which have already been published were submitted before the author changed her name, and as such the name "Alexander Jones" shall be treated as equivalent. Exceptions to sole authorship of material presented here are as follows:

Research presented in Chapter 4:

This research was conducted at the University of Waterloo, the printed circuit board and mechanical relay used to measure the resistance of the film was initially designed by Manfred Kao (M.ASc), Dr. Kissan Mistry, and Ahmed Shahin under the dual supervision of Dr. Kevin Musselman and Dr. Mustafa Yavuz. The programming of the control system and analysis of the data was performed by Olivia Marchione, under the supervision of Dr. Kevin Musselman.

A. Jones, K. Mistry, M. Kao, A. Shahin, M. Yavuz, and K. P. Musselman, "In-situ spatial and temporal electrical characterization of ZnO thin films deposited by atmospheric pressure chemical vapour deposition on flexible polymer substrates," *Scientific Reports 2020 10:1*, vol. 10, no. 1, pp. 1–10, Nov. 2020, doi: 10.1038/s41598-020-76993-4.

Research Presented in Chapter 6:

This research was conducted at the University of Waterloo in collaboration between Olivia Marchione and Dr. Kissan Mistry, under the supervision of Dr. Kevin Musselman. Dr. Kissan Mistry was responsible for the initial selection of the reflectance spectrometer used, and

developed the grid-search algorithm. Olivia Marchione was the sole developer of all machine-learning related code, and the control system, and worked together with Dr. Kissan Mistry to program the in-situ measurement system.

K. Mistry, A. Jones, M. Kao, T. W.-K. Yeow, M. Yavuz, and K. P. Musselman, " In-situ observation of nucleation and property evolution in films grown with an atmospheric pressure spatial atomic layer deposition system ," *Nano Express*, vol. 1, no. 1, p. 010045, Jun. 2020, doi: 10.1088/2632-959x/ab976c.

Research presented in Chapter 7:

This research was conducted at the University of Waterloo by Olivia Marchione under the supervision of Dr. Kevin Musselman. Olivia Marchione designed this experiment and performed all related programming tasks as well as taking ellipsometry measurements and performing initial model fitting. Ambuj Ranjan assisted with collecting ellipsometry measurements as part of an undergraduate research project. Dr. Farman Ullah provided support by measuring the film roughness with an atomic force microscope, and Hatameh Asgarimoghaddam assisted with some of the absorbance spectroscopy measurements.

This research is in preparation to be published.

Research presented in Chapter 8:

This research was conducted at the University of Waterloo by Olivia Marchione under the supervision of Dr. Kevin Musselman. Dr. Abdullah Alshehri was responsible for the deposition of the perovskite layers and gradient alumina barrier, with the assistance of Dr. Khaled Ibrahim and Olivia Marchione. Hatameh Asgarimoghaddam was responsible for the deposition of the calcium layers, and conducting the water vapour transmission rate calculations. Olivia

Marchione was the sole programmer of the computer-vision system, responsible for it from conception to final implementation.

A. H. Alshehri *et al.*, “Nanoscale Film Thickness Gradients Printed in Open Air by Spatially Varying Chemical Vapor Deposition,” *Adv Funct Mater*, vol. 31, no. 31, p. 2103271, Aug. 2021, doi: 10.1002/ADFM.202103271.

The calcium film paper is in preparation to be published.

Abstract

The deposition and design of nanometre-scale oxide films is an integral component of the ongoing nanomaterial revolution, from cell phones, to batteries, to photovoltaics. Atmospheric-pressure spatial atomic layer deposition (AP-SALD) and chemical vapour deposition (AP-CVD) are two techniques that show great promise for commercialization, due to their speed and the vast array of materials and stoichiometries they can produce. However, this flexibility comes at the cost of complexity; the presence of oxygen during film deposition induces defects which may have advantageous or deleterious effects dependent on the application of the film. Machine learning is a statistical technique that allows us to make sense of such complex systems of interaction, without the need for expensive ab-initio simulations. Through this work, I demonstrate our capacity to deposit entirely new materials with our lab-scale AP-SALD/CVD system, and develop several characterization methods that will permit us to leverage the flexibility of that system to maximum effect. I developed a system for measuring the resistance of our films in-situ, demonstrating the effects of atmospheric oxygen on film properties, as well as implementing machine learning into our in-situ reflectance system providing accurate real-time measurements of film thickness and band gap. Next, I designed a gaussian process regression tool to assist researchers in finding accurate optical model parameters with our spectroscopic ellipsometer, much more quickly than previously possible. Lastly, I implemented a basic computer vision tool for tracking the degradation of perovskite and calcium thin films in real-time. To my knowledge, this work represents the first time that machine learning has been leveraged to improve the deposition of films by AP-SALD and to enhance the characterization of their properties and performance.

Acknowledgements

I would like to acknowledge the support and supervision of Dr. Kevin Musselman, who has been incredibly understanding of the myriad challenges I have faced over the course of completing this degree, and pass on my deepest and most sincere thanks.

Table of Contents

Examining Committee Membership ii

Authors Declaration iii

Statement of Contributions iv

Abstract vii

Acknowledgements viii

List of Figures xii

List of Tables xvii

List of Abbreviations xviii

1 Introduction ..... - 1 -

1.1 Thesis Objective ..... - 2 -

1.2 Thesis Outline ..... - 4 -

Part A - AP-SALD and In-Situ Data Collection ..... - 6 -

2 Literature Review: Atmospheric-Pressure Spatial Atomic Layer Deposition ..... - 7 -

3 New Precursor Selection and First Thermal AP-SALD Copper films ..... - 18 -

3.1 Introduction ..... - 18 -

3.2 Experimental Methods and Materials ..... - 19 -

3.3 Results and Discussion ..... - 20 -

3.4 Conclusions ..... - 28 -

4 In-Situ Electrical Characterization ..... - 29 -

4.1 Introduction ..... - 29 -

4.2 Experimental Methods and Materials ..... - 30 -

4.2.1 PCB and Control System ..... - 30 -

4.2.2 Film Deposition and Characterization ..... - 32 -

4.3 Results and Discussion ..... - 32 -

4.4 Conclusion ..... - 40 -

Part B - Machine Learning in Materials Science ..... - 41 -

5 Literature Review: Data Collection for Machine Learning in Materials Science ..... - 42 -

5.1 Materials Science and Characterization ..... - 44 -

5.1.1 Reflectance Spectroscopy - 45 -
5.1.2 Ellipsometry - 49 -
5.1.3 Optical Models for Reflectance and Ellipsometry - 57 -
5.1.4 Absorbance Spectroscopy - 64 -
5.2 Machine Learning - 65 -
5.2.1 Gaussian Process Regression - 78 -
5.2.2 Convolutional Neural Networks - 80 -
5.2.3 Material Science with ML - 82 -
6 Neural Network Assisted In-Situ Optical Characterization and Feedback Control - 88 -
6.1 Introduction - 88 -
6.2 Experimental Methods and Materials - 91 -
6.2.1 Film Deposition and Characterization Techniques - 91 -
6.2.2 Reflectance Measurement Methods - 92 -
6.2.3 Machine Learning Methods - 93 -
6.3 Results and Discussion - 96 -
6.3.1 Grid Search Algorithm - 96 -
6.3.2 Feed-Forward Artificial Neural Network - 99 -
6.3.3 Gaussian Process Regression - 104 -
6.3.4 Automatic Feedback for AP-SALD System - 107 -
6.4 Conclusion - 110 -
7 ML-Empowered Ellipsometry - 111 -
7.1 Introduction - 111 -
7.2 Experimental Methods and Materials - 114 -
7.2.1 AP-SALD Depositions - 114 -
7.2.2 Sample Preparation - 115 -
7.2.3 Data Preparation - 117 -
7.2.4 Machine Learning Methods - 119 -
7.3 Results and Discussion - 124 -
7.4 Conclusions - 134 -

8 Computer Vision for Perovskite and Calcium Film Degradation Detection ..... - 136 -
8.1 Introduction ..... - 136 -
8.2 Experimental Methods and Materials ..... - 138 -
8.2.1 Sample Preparation ..... - 138 -
8.2.2 Film Degradation Detection ..... - 140 -
8.2.3 Machine Learning Methods ..... - 143 -
8.3 Results and Discussion ..... - 144 -
8.4 Conclusions ..... - 150 -
9 Conclusions and Next Stages ..... - 151 -
9.1 Machine-Learning Integrated Control System ..... - 153 -
9.2 Discovery of New ε-Near Zero Materials ..... - 156 -
Letters of Copyright Permission ..... - 160 -
References ..... - 167 -

List of Figures

Figure 2.1: Chemical Vapour Deposition process in vacuum, precursor compounds are introduced simultaneously and allowed to react on the substrate. ..... - 10 -

Figure 2.2: Vacuum or conventional Atomic Layer Deposition, precursors are allowed to adsorb onto the surface sequentially, limiting the rate of reaction. ..... - 11 -

Figure 2.3: Oscillating-stage style Atmospheric-Pressure Spatial Atomic Layer Deposition. Precursors adsorb to the substrate sequentially, separated in space rather than time. ..... - 12 -

Figure 2.4: Labelled schematic of the lab-scale AP-SALD system in use at the University of Waterloo ..... - 13 -

Figure 2.5: How liquid and solid precursors are mixed with inert nitrogen gas to be delivered to the AP-SALD reactor. ..... - 14 -

Figure 3.1: Two films deposited using CupraSelect™ as the copper precursor. a) and c) show a film deposited at 200ºC on a black and white background respectively. b) and d) show a film deposited at 125ºC on the same backgrounds. On the black background the metallic copper is more easily visible, whereas the copper oxides are more easily seen on paper. ..... - 21 -

Figure 3.2: Four films deposited using $Cu(hcaf)_2$ at 200ºC with various co-reactants, metallic copper highlighted in parts a) and b). a) shows a film deposited with $Cu(hcaf)_2$ alone, with no co-reactant. b) shows a film deposited with 10 SCCM of DEZ added to the reactor. c) shows a film deposited with 10 SCCM of pyridine added as a reducing agent, and d) shows a film with 20 SCCM of deionized water added, which acted as an oxidant. ..... - 23 -

Figure 3.3: X-ray diffraction scan of three copper-containing samples deposited at 200ºC. In blue is a film deposited with the $Cu(hcaf)_2$ precursor alone, in green the same conditions were used but diethylzinc was tested as a reducing agent, and in red is a film deposited with CupraSelect™ and no co-reactant. ..... - 24 -

Figure 3.4: Electron microscope image of the 200ºC film deposited with no co-reactant. Energy-dispersive x-ray spectroscopy scans show there are large gaps in the film based on the intensity of the copper peak in selected area 1 vs. 2. ..... - 27 -

Figure 4.1: PCB design for in-situ resistance measurements – Source: Olivia Marchione [https://www.nature.com/articles/s41598-020-76993-4] ..... - 30 -

Figure 4.2: In-situ electrical measurement system. A flexible polyimide PCB substrate is placed on the oscillating substrate stage of the AP-SALD/SCVD system and electrical resistance is measured at 20 trace pairs, shown in red,

across the substrate throughout the deposition of a thin film. measurements – Source: Olivia Marchione [https://www.nature.com/articles/s41598-020-76993-4] ..... - 31 -
Figure 4.3: (a) All resistances measured throughout the deposition of a ZnO film on the polymer substrate at 175°C. A selection of resistances measured across the polymer substrate during the deposition of a ZnO film at (b) 100°C, (c) 125°C, (d) 150°C, and (e) 175ºC measurements – Source: Olivia Marchione [https://www.nature.com/articles/s41598-020-76993-4] ..... - 34 -
Figure 4.4: Change in resistance over time at the centre of the film for depositions at (a) 150°C on a linear plot and (b) all deposition temperatures on a log-linear plot. Three distinct growth regions are highlighted in (a). measurements – Source: Olivia Marchione [https://www.nature.com/articles/s41598-020-76993-4] ..... - 35 -
Figure 4.5: AFM images of (a) a bare silicon wafer and silicon wafers with (b) 5 oscillations and (c) 30 oscillations of ZnO deposition at 150°C. Island formation and coalescence is observed. measurements – Source: Olivia Marchione [https://www.nature.com/articles/s41598-020-76993-4] ..... - 37 -
Figure 4.6: Change in resistance of trace pair 10 from the 150°C film as it is heated on a hot plate from room temperature to 180°C. measurements – Source: Olivia Marchione [https://www.nature.com/articles/s41598-020-76993-4] ..... - 39 -
Figure 4.7: Resistance across the PCB's measured as deposited and at room temperature, films deposited at a) 100°C, b) 125°C, c) 150°C, and d) 175°C measurements – Source: Olivia Marchione [https://www.nature.com/articles/s41598-020-76993-4] ..... - 40 -
Figure 5.1: Schematic of the in-situ reflectance spectrometer, and three-layer Fresnel model. The oscillating stage is programmed to pause outside the reactor head of the AP-SALD system for the duration of the measurement, to ensure that no errors are introduced by the motion. ..... - 45 -
Figure 5.2: Reflectance spectrum of a ZnO film deposited at 150ºC and measured from 250-800nm (blue) compared to the best-fit spectrum of a film with a known thickness and complex refractive index (red). ..... - 47 -
Figure 5.3:Graphical diagram of; a) unpolarized, b) linearly polarized, c) circularly polarized, and d) elliptically polarized light. ..... - 50 -
Figure 5.4: Two sets of 25 measurements taken from different locations on the same ZnO film. In example a) all model parameters were held constant, with only thickness varying. In example b) four model parameters were allowed to vary between measurements, resulting in a significantly lower value of FD. ..... - 53 -

Figure 5.5: Idealized schematic of Fit Difference vs one model parameter, showing local minima. ..... - 54 -
Figure 5.6: Parameterization of an alumina film deposited using the lab AP-SALD to the Cauchy optical model within the FS-1 ellipsometer. ..... - 59 -
Figure 5.7: Refractive index of nitrogen-doped AZO film calculated using the Tauc-Lorentz optical model. ..... - 62 -
Figure 5.8: Roughness and its approximation within an optical model. ..... - 64 -
Figure 5.9: Two- and Three-dimensional representations of common issues arising from poor hyperparameter selection during gradient descent. ..... - 68 -
Figure 5.10: Illustrations of the three types of ML problem, a) Classification, b) Clustering, and c) Regression .. - 70 -
Figure 5.11: Four common nonlinear activation functions ..... - 71 -
Figure 5.12:Generalized data flow within a neural network ..... - 73 -
Figure 5.13:Predicted curve with varying data density ..... - 75 -
Figure 5.14: The effect of adding extra dimensions on previously non linearly separable data through the kernel trick.: ..... - 76 -
Figure 5.15: Linear regression (left) vs Gaussian process regression (right) over the same arbitrary data points.. - 78 -
Figure 5.16: Convoluting an input and filter to find a feature map of a simple binary matrix ..... - 80 -
Figure 5.17: Three most common pooling techniques ..... - 81 -
Figure 6.1: Workflow schematic between the reflectometer, search algorithm, and film database. ..... - 91 -
Figure 6.2: Estimated film thicknesses based on a grid search around an assumed growth rate of 1 nm/oscillation, either a) unconstrained, or b) constrained to within ±5nm for the first 30nm, and then to within ±50nm. ..... - 96 -
Figure 6.3: a) Reflectance spectrum changes with film growth up to 30 oscillations, b) changes in measured thickness and goodness of fit, and c) evolution of band gap and real refractive index over the same interval – Source: Kissan Mistry [https://iopscience.iop.org/article/10.1088/2632-959X/ab976c] ..... - 98 -
Figure 6.4: Nonlinearity after the nucleation phase still visible, even with the most advanced grid-search algorithm we developed. ..... - 99 -
Figure 6.5:Example of fitted film thickness deposited at 150ºC using first generation ANN, with significantly smoother fitted thickness values above 60 nm compared to the model with no ML-assisted fitting ..... - 102 -
Figure 6.6: Plot of film thickness deposited at 150ºC vs oscillation number, comparing ML-assisted fitting and basic grid-search methods at high oscillations – image courtesy of Kissan Mistry ..... - 103 -

Figure 6.7: a) Actual (blue) vs predicted (yellow) values for reflectance data during training of the v2 GPR model, noise reduction at low thicknesses is visible. b) Training histogram for the v2 GPR model, iteration with lowest observed MSE indicated in yellow ..... - 105 -
Figure 6.8: Fitted Thickness of ZnO film deposited at 150ºC vs Oscillation Number using the second-generation GPR algorithm. ..... - 106 -
Figure 6.9: Changes in ZnO growth rate due to increased DEZ-laden nitrogen flow ratio through the bubbler, ratio of DEZ:Carrier indicated for each section – image courtesy of Kissan Mistry ..... - 107 -
Figure 6.10: Response time of a ZnO film deposition at 150ºC controlled by our in-situ PI algorithm ..... - 110 -
Figure 7.1: Schematic of how samples are prepared for inclusion within the training dataset. a) AP-CVD-mode deposition, with limited mixing occurring below the reactor head. b) A similar deposition occurring in AP-SALD-mode, where each precursor adsorbs fully before reacting with the next. c) How we can leverage the precursor mixing of an AP-CVD deposition to produce gradient films, where differing mixing rates lead to differing deposition rates along the width of the substrate. d) AP-SALD cannot produce gradient films, as the rate of reaction is limited by the complete adsorption of precursor onto the substrate. e) and f) AP-CVD gradient and AP-SALD uniform film produced at the same temperature of 200℃ using 250 oscillations. g) Schematic of how each film sample is broken down and measured to provide hundreds of measurements per film. ..... - 116 -
Figure 7.2: ML model workflow, showing how the output of the model is iteratively improved using the ellipsometers inbuilt LM fitting before it's used to train the next iteration. ..... - 118 -
Figure 7.3: Division of the total measurements taken for this work, the bulk were reserved for validation and testing, but the differences between various iterations of the model are shown. ..... - 119 -
Figure 7.4: Illustration of the common causes of graded optical properties in AP-SALD films. ..... - 126 -
Figure 7.5: Average fit difference of all films in each iteration ..... - 128 -
Figure 7.6: Raw ellipsometry data and average fit difference of two films used in this study. On the left is a 50nm ZnO film produced under uniform ALD conditions, and on the right is a Cu:N:ZnO film produced under gradient CVD conditions with a thickness ranging from ~250-300nm. ..... - 129 -
Figure 7.7: AFM image procured from three different 10x10µm areas of an AZO film. a) Edge of the deposition area, b) centre of the deposition area, and c) outside the deposition area, where precursors flowed out from below the reactor. ..... - 131 -

Figure 7.8: Tauc plots of three samples of varying composition and thickness, from left-to-right we have Cu_2O, ZnO and AZO films with the band gap indicated in blue. ..... - 133 -
Figure 8.1: Film stacks used in phase one and two of this work, showing a) perovskite films with gradient alumina and b) calcium films with uniform alumina and added background layer. ..... - 138 -
Figure 8.2: The same perovskite film with alumina barrier layer pictured after 0, 48, and 92 hours at 60% relative humidity and room temperature. ..... - 140 -
Figure 8.3: Three calcium films imaged at 1 and 72 hours. Varying ambient lighting conditions resulted in a much less reliable colour difference between intact and degraded films, so a red background layer was added for ease of identification of degraded and intact areas. ..... - 141 -
Figure 8.4: Schematic of the convolutional neural network developed for phase one of this work. ..... - 143 -
Figure 8.5: Final results of the perovskite degradation CNN, it can be seen that as time progresses the degraded squares marked in red show the same progression as the unedited film images. ..... - 146 -
Figure 8.6: Processed vs original image of a calcium film trained with no added background. Green denotes areas classified as degraded. Notable features indicated, 1) is an angled line that was well captured by the algorithm, 2) is an open area of degraded film not captured by the algorithm, and 3) is an area of pinhole-defects that were also not well-resolved. ..... - 147 -
Figure 8.7: Re-trained calcium WVTR network using a red background to enhance visual contrast between degraded and intact films. ..... - 149 -
Figure 9.1: Metal-oxide semiconductor field effect transistor (top) vs. ENZ optical transistor (bottom) ..... - 158 -

List of Tables

TABLE I: FILM PROPERTIES AT SELECTED DEPOSITION TEMPERATURES ..... - 38 -

TABLE II: MATERIAL SCIENCE USING MACHINE LEARNING AND VARIOUS DATA SOURCES ..... - 44 -

TABLE III: MATERIAL PROPERTY PREDICTION AND CHARACTERIZATION USING MACHINE LEARNING ..... - 85 -

TABLE IV: COMPARISON OF THE PERFORMANCE OF FEED-FORWARD ANN'S OF VARYING WIDTH, TRAINED USING LEVENBERG-MARQUARDT, BAYESIAN REGRESSION, AND SCALED CONJUGATE GRADIENT ALGORITHMS. ..... - 101 -

TABLE V: BAND GAP AND INTERBAND TRANSMISSION VALUES FOR MATERIALS USED IN THIS STUDY ..... - 122 -

TABLE VI: DRUDE PARAMETERS, PLASMA FREQUENCY AND DAMPING FOR ALL RELEVANT MATERIALS ..... - 123 -

TABLE VII: R^2 AND RMSE VALUES OF FOUR ML MODELS COMPARED, BEST RESULTS HIGHLIGHTED IN BOLD ..... - 125 -

TABLE VIII: RMSE COMPARISON OF GPR MODELS PRODUCED WITH COMBINATIONS OF DATA ABOUT THE DEPOSITION PROCESS, BEST RESULTS SHOWN IN BOLD. ..... - 128 -

TABLE IX: TRAINING, VALIDATION, AND TESTING PERFORMANCE OF FOUR CNN ALGORITHMS OF INCREASING DEPTH ..... - 145 -

List of Abbreviations

AFM	Atomic Force Microscopy
ALD	Atomic Layer Deposition
ANN	Artificial Neural Network
AP-CVD	Atmospheric Pressure Chemical Vapour Deposition
AP-SALD	Atmospheric Pressure Spatial Atomic Layer Deposition
AZO	Aluminium-Doped Zinc Oxide
BR	Bayesian Regression
CVD	Chemical Vapour Deposition
CNN	Convolutional Neural Network
DEZ	Diethylzinc
DFT	Density Functional Theory
DNN	Deep Neural Network
FD	Fit Difference
GAN	Generative Adversarial network
GPR	Gaussian Process Regression
KNN	K-Nearest-Neighbours
LDA	Linear Discriminant Analysis
LM	Levenberg-Marquardt
LSTM	Long Short-Term Memory
ML	Machine Learning
MOSFET	Metal Oxide Semiconductor Field Effect Transistor
NMF	Nonnegative Matrix Factorization

NN	Neural Network
PAC	Programmable Automation Controller
PCA	Principal Component Analysis
PCB	Printed Circuit Board
ReLU	Rectified Linear Units
RNN	Recurrent Neural Network
RHEED	Reflection High-Energy Electron Diffraction
SALD	Spatial Atomic Layer Deposition
SCG	Scaled Conjugate Gradient
SVM	Support Vector Machine
TDMASn	Tetrakis(dimethylamino)tin(IV)
TMA	Trimethyl Aluminium
WVTR	Water Vapour Transmission Rate
XOR	Exclusive-Or
XRD	X-Ray Diffraction

1 Introduction

Nanomaterials are being used more and more every year, in devices we use every day. The screens of our cell phones and televisions, the transistors in our computers, diodes, batteries, photovoltaics, nanomaterials are everywhere. In order to meet increasing demands for energy efficiency and high-performance it has become necessary to design nanomaterials with specific properties for individual applications. To keep such specificity of design within the realm of commercial viability, rapid deposition techniques are needed. Two such techniques are atmospheric-pressure spatial atomic layer deposition (AP-SALD) and atmospheric-pressure chemical vapour deposition (AP-CVD). Because these are performed in atmosphere, there is no time-consuming vacuum purge step – particularly important in the case of AP-SALD, as regular atomic layer deposition (ALD) needs to be purged multiple times throughout the deposition. To fully leverage the flexibility of this system, it would either be necessary to fully model the interactions between all the various compounds used to determine the best conditions for producing a given film, or use statistical techniques such as machine learning (ML). Simulations of atomic interactions can be very time-consuming, and rely inherently upon assumptions and simplifications that may result in a loss of prediction accuracy, but also allow us to fully understand the interactions at play within our system. ML can be built on a foundation of real data, collected from real films and so explicitly captures dynamics we may fail to account for in a simulation, but is also a "black box" where data goes in and results come out – without necessarily enhancing our understanding of the relationships therein. However, we are interested here in the *applications* of the films produced using AP-SALD/CVD, and so it is more important that we be able to accurately predict the properties of real films than understand how those properties arose.

This manuscript will outline my research towards using ML tools, specifically Gaussian Process Regression (GPR) and Convolutional Neural Networks (CNN), to predict the material properties of thin metal chalcogenide films produced using our AP-SALD/CVD system. These films can then be tailored for use in specific devices, such as photovoltaics, anti-viral coatings, or wearable electronics. The first phase of this research involved the improvement of our lab's AP-SALD system, and development of in-situ measurement tools to help ensure the reliability of the system when it came time to create the ML training dataset. The second phase of my research involved the development and initial utilization of this dataset, creating an ML model for use with a Film Sense FS-10X ellipsometer to rapidly and precisely fit films of widely varying composition. Some of these films were then tested within devices, specifically passivation layers for use in perovskite photovoltaics. The effectiveness of these passivation layers was determined visually using a CNN model which can detect degradation of the subject film when using either a real perovskite or test material.

1.1 Thesis Objective

At the end of this thesis I will have answered the question of how we can use our AP-SALD system and machine learning techniques to accelerate the advance of materials science at the nanometre scale, with a focus on characterization of novel thin films. There is a large degree of variability between AP-SALD systems, as there are few commercially-produced systems available at this time. This means that the correlation between the parameters used to control our system and the properties of the films produced by said system is poorly understood. Even just using our system there are fourteen process parameters which directly affect the film, discounting completely the library of chemical precursors we can select. The lack of data and complex data-space makes the AP-SALD a challenging process to optimize; fortunately ML

techniques have been applied to similar problems [1], [2]. ML has shown to be an incredibly useful tool for materials scientists as well as process engineers, like when screening large sets of data either from databases or experiments to find novel new materials [3]–[8]. Some researchers have applied machine learning techniques to conventional atomic layer deposition (ALD) processes [9], and some used ML in conjunction with simulation techniques like Density Functional Theory (DFT) to optimize the ALD cycles [10], [11]. In both cases conventional ALD was used, which is a vacuum technique and so is unaffected by the presence of oxygen annealing during deposition (unlike AP-SALD), which changes the role of multiple process parameters but requires significantly longer to produce films of equal thickness compared to our technique. This means that films produced with AP-SALD are better suited for rapidly exploring large composition and parameter data-spaces. Both the above works limited themselves to a single combination of precursor chemicals, with fewer alterations to the process parameters.

This research will focus on the methods used to generate and relate large sets of data in a system of zinc, aluminium, copper, oxygen and nitrogen precursors. These films are produced with varying amounts of the precursor chemicals, including binary oxide films, at various deposition conditions. We already have the capability to measure the thickness and optical properties (refractive index, bandgap) of the film during and after production, and techniques for measuring resistivity directly were developed but eventually discarded for reasons that will be explained in the relevant chapters. Ellipsometry and absorbance measurements proved to be critical in the study of these films, providing the thickness, band gap, roughness, and even information on the resistivity and plasma frequency of the films. Herein lay the first challenge towards developing an accurate ML program, developing a dataset comprehensive and accurate enough to build an accurate ML model. This stage was completed in parts, with the initial film

preparation and measurements consuming nearly an entire year. Once the initial data had been collected, however, initial versions of each network could be created – speeding up the creation of the subsequent iteration through online learning. The first ML model created was designed to provide initial guesses for the modelling parameters used by the ellipsometer. This sped up the collection of ellipsometry data massively, as each sample had to be fit only once, not iteratively to reduce fitting error to acceptable levels. By using an existing optical model, we were able to leverage the underlying assumptions behind the GPR model, that a function exists which describes the relation between input and output parameters. This model also accounts for the natural variability that is inherent to any measurement technique that relies upon solving a complex inverse problem, such as ellipsometry.

Machine learning techniques are not just useful for *producing* films, however, they can also be of great use when testing device and film stacks. For example, ALD alumina films have been used as passivation layers before, protecting humidity- and oxygen-sensitive films like perovskite photovoltaics. In this research we tested the humidity-resistance of multiple passivation layers, using either real perovskite film stacks or an evaporated calcium layer as the water-sensing layer. By using a CNN, the degradation of these films was able to be monitored in real-time, providing much more precise measurements of the water vapour transmission rate (WVTR) than is possible by eye alone.

1.2 Thesis Outline

Part A of this manuscript is a review of the literature relating to AP-SALD, the selection of a new precursor chemical and subsequent deposition of some of the first thermal AP-SALD copper/oxide films, and the development of an in-situ resistance sensing technique initially

conceived to increase the amount of data collected from each film deposited. These techniques were later discarded in favour of more advanced tools, but they form part of a potential feedback system to better control the AP-SALD process. Part B explains the work I completed in order to refine our in-situ reflectance measurement system, produce the ML model designed to speed up the process of collecting data from the ellipsometer, and the ways in which basic computer vision can be used to automate time-intensive material degradation tests. Both of these sections will be preceded with a targeted review of the subject, including relevant background as well as contemporary research, and the final section of my report will be a projection of how this work can be used in future research.

Part A - AP-SALD and In-Situ Data Collection

The first main aspect of this project was the development of various tools for creating and characterizing films using our AP-SALD system. First, I will review the AP-SALD process, and some of its uses as demonstrated by us and our collaborators. Then I will discuss my efforts to expand our library of materials to include copper and its oxides in response to the COVID-19 pandemic, and to develop a method of measuring film resistance in-situ.

2 Literature Review: Atmospheric-Pressure Spatial Atomic Layer Deposition

Atomic layer deposition is a variant of the chemical vapour deposition (CVD) process for creating nanometre-thick films, usually metal chalcogens such as zinc oxide. All of the techniques presented here involve reactions between two or more chemical precursors diluted in an inert carrier gas such as nitrogen. Each specific compound is formed by a unique and complex combination of chemical reactions, which can make it challenging to understand exactly what is occurring during the deposition [12]. Which reactions occur is based on the thermodynamics of those reactions, and competing reactions can occur within the same deposition if the conditions are not selected well. Of these reactions, the following are most relevant to the compounds we most often produce in our lab:

- **Pyrolysis**: Decomposition of a molecule due to thermal energy breaking the chemical bonds, without interacting chemically with another compound. This can be useful for breaking down a precursor in preparation for a reaction, but can also lead to degradation of the deposited film.
- **Hydrolysis**: A decomposition reaction like pyrolysis, but initiated by a reaction between a compound and water. Water is a commonly-used source of oxygen for depositing oxide films, and so can perform multiple roles simultaneously.
- **Reduction and Oxidation**: These two kinds of reactions always occur together, involving an exchange of electrons between reactants, and are often referred to by the combined term "Redox Reaction". The reactant which gains electrons is said to be reduced, and the compound which loses electrons is oxidized, but oxygen does not have to be involved for this

reaction to occur. Redox reactions are particularly important when one of the elements in the reaction has multiple oxidation states; copper is an example of a material that can form multiple oxides or a pure metal depending on its electronegativity.

- **Disproportionation**: Often a kind of redox reaction where a reactant is simultaneously oxidized and reduced, but refers to any reaction where one compound breaks down into two different products. For example, in some cases we used hydrogen peroxide instead of water as the oxygen source, which breaks down into water and oxygen gas – increasing the concentration of oxygen in the reaction volume.

Which of these reaction types dominates the process, and therefore the stoichiometry of the final film, can be tuned through the addition of energy in the form of heat, plasma, or ultraviolet light. We can predict which compounds will be produced depending on the thermodynamics of the reaction involved, specifically whether the free energy change of the reaction (G_r) is negative, which is dependent on the Gibbs free energy of each reactant and product (G_f) through the simple equation:

$$G_r = \sum G_{f,products} - \sum G_{f,reactants} \quad (1)$$

G_f depends not only on the properties of the compound itself, but also on the temperature, pressure, and molar concentration of each material. It is through these last three parameters that we can exert control over the reaction taking place, however, discovering new reactants with more favourable properties is outside the scope of this research. Many different classes of compounds have been used as precursors over the years, but modern precursors are often highly volatile and reactive metallo-organic compounds. Due to their volatility, these precursors can often be used at significantly lower temperatures than previous chemicals, but also necessitates

careful separation from the environment. Some specific risks associated with metallo-organic precursors are discussed in depth in Chapter 3, where I describe the process we use to select a new precursor. The reaction dynamics described above apply to all vapour-phase reactions and deposition techniques including our AP-SALD system, which is an evolution of conventional ALD designed for greater commercial applicability. CVD, ALD, and AP-SALD are three very similar vapour-phase deposition techniques, but each has its unique advantages and disadvantages.

- **CVD**, illustrated in Figure 2.1, is a process where the precursor chemicals are injected simultaneously into a vacuum chamber and allowed to react on the surface of the substrate. This technique offers high growth rates, as the reaction rate is not limited by the amount of precursor adsorbed to the surface. However, the films have difficulty adhering to high aspect ratio features, and can have defects such as pores in the film. More critically, as the reaction in CVD occurs much more rapidly, we cannot obtain as fine control of the thickness of the film, especially for very thin films. Because the precursors mix throughout the deposition, there is no reason to purge the vacuum chamber until the deposition is complete. Given that it takes a significant amount of time to reach the level of vacuum needed for thin film deposition processes, the reduced number of purge cycles has been a boon for the commercialization of CVD.

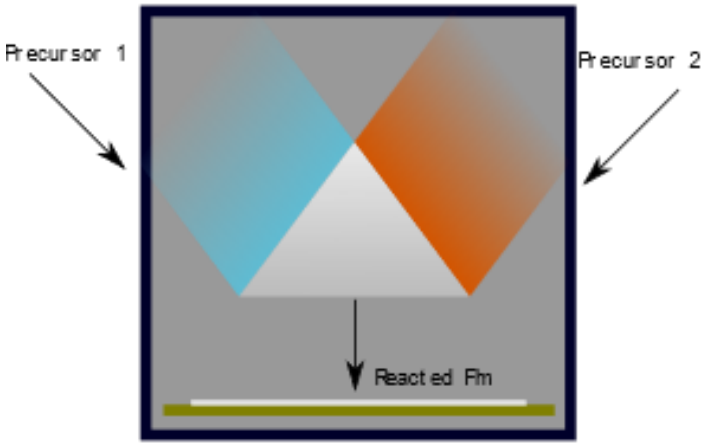

Figure 2.1: Chemical Vapour Deposition process in vacuum, precursor compounds are introduced simultaneously and allowed to react on the substrate.

- **ALD**, shown in Figure 2.2, differs from CVD in that each of the precursor chemicals is deposited onto the substrate separately, and each deposition is separated by a purge step. These purge steps can take a long time, as all excess precursor must be removed from the vacuum chamber. Injecting the precursors one at a time allows a thin layer of precursor to adhere to the entire substrate. Since the precursor which did not adhere to the substrate is removed, the reaction is limited by how much can adsorb onto the substrate. This is called a self-limiting reaction. It is this self-limiting nature that gives ALD much finer control over the final film thickness, as each deposited layer is one atom thick. ALD is a much slower process than CVD due to the lengthy purge cycles, but produces films with few defects that can adhere to high aspect-ratio features.

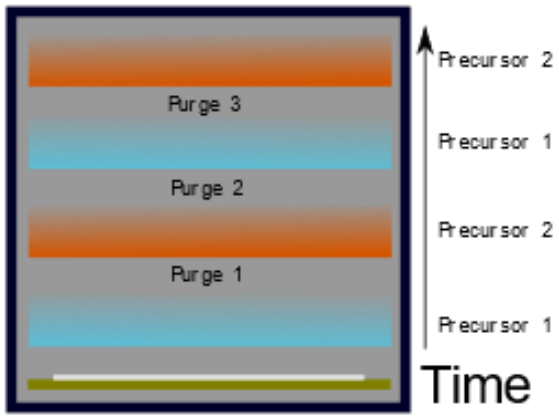

Figure 2.2: Vacuum or conventional Atomic Layer Deposition, precursors are allowed to adsorb onto the surface sequentially, limiting the rate of reaction.

- **AP-SALD**, shown in Figure 2.3, is a proposed solution to the issue of long purge times – separating the precursor flows in space rather than time [13]. Instead of long purge cycles between each precursor addition, all precursors are directed to flow onto the substrate at the same time by a specially designed reactor head. The substrate is then made to oscillate beneath this reactor head, sequentially exposing each part of the substrate to the precursors. To prevent mixing of the precursors above the substrate, the flows are separated by inert gas curtains and the excess precursor is removed using a vacuum. This technique can be performed in atmospheric conditions, as the inert gas separates the precursors from the environment and one another before they react on the substrate [14]. In the close-up portion of Figure 2.3 we can see how the successive layers of each precursor react with one another, building up multiple ALD cycles in a single oscillation of the stage. Conventionally, AP-SALD is operated in the same self-limiting manner as conventional ALD, however it is possible to weaken the separation between the precursor flows or increase the total flow, such that there is some mixing of the precursors in the gas phase and the system operates as an atmospheric-pressure CVD system instead.

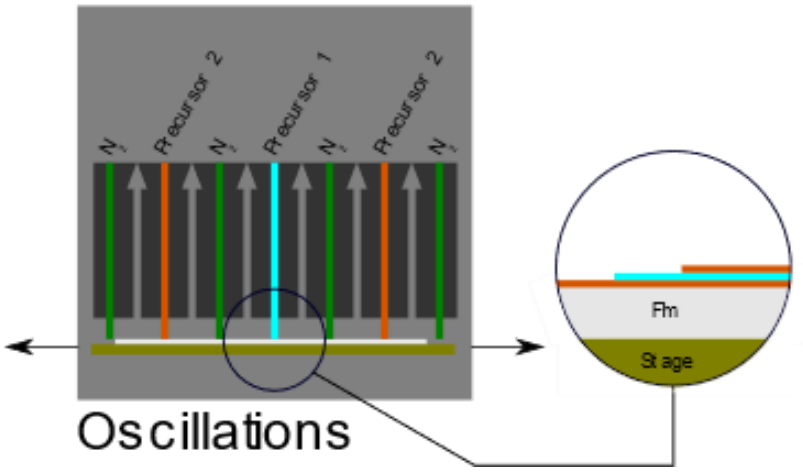

Figure 2.3: Oscillating-stage style Atmospheric-Pressure Spatial Atomic Layer Deposition. Precursors adsorb to the substrate sequentially, separated in space rather than time.

The quality and morphology of films created using AP-SALD depends heavily on both process parameters set by the user, such as the temperature of the substrate and the flow rates of each precursor vapour, as well as local conditions such as humidity and the number of depositions performed that day. As the system is not purged of the precursor compounds it is possible for the concentration of the reactive species to increase over time within the tubes delivering them to the reactor, and therefore affecting the growth rate of the film. Even when creating an undoped, homogenous film, it can be difficult to select the correct process parameters which will result in a film with the desired properties.

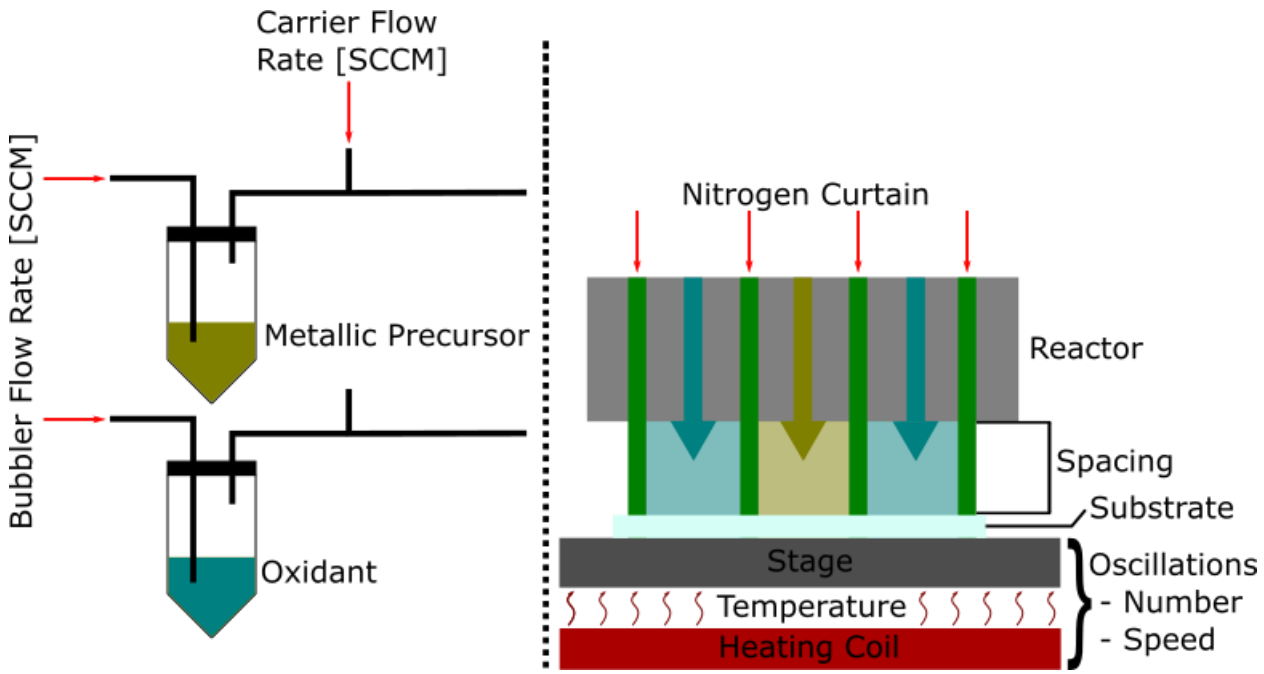

Figure 2.4: Labelled schematic of the lab-scale AP-SALD system in use at the University of Waterloo

Stage temperature, reactor-substrate spacing, oscillation speed, number of oscillations, nitrogen curtain strength, and the total of carrier and bubbler flow rates are selected for each deposition in our system, illustrated in Figure 2.4. These seven parameters are constant for all of the lines that deliver the different chemical precursors at each point in time for the sake of producing uniform depositions, but the precursor parameters can vary between each chemical precursor line. The two parameters that we can vary between precursor lines are the percentage of carrier flow flowing through the bubbler, and which channels of the reactor the precursor-laden flow passes through. A minimum of two precursors is required, so more than ten parameters can vary between depositions. Added to the time-dependent variability in growth rates, there is a very large data-space that must be understood to attain proper process control.

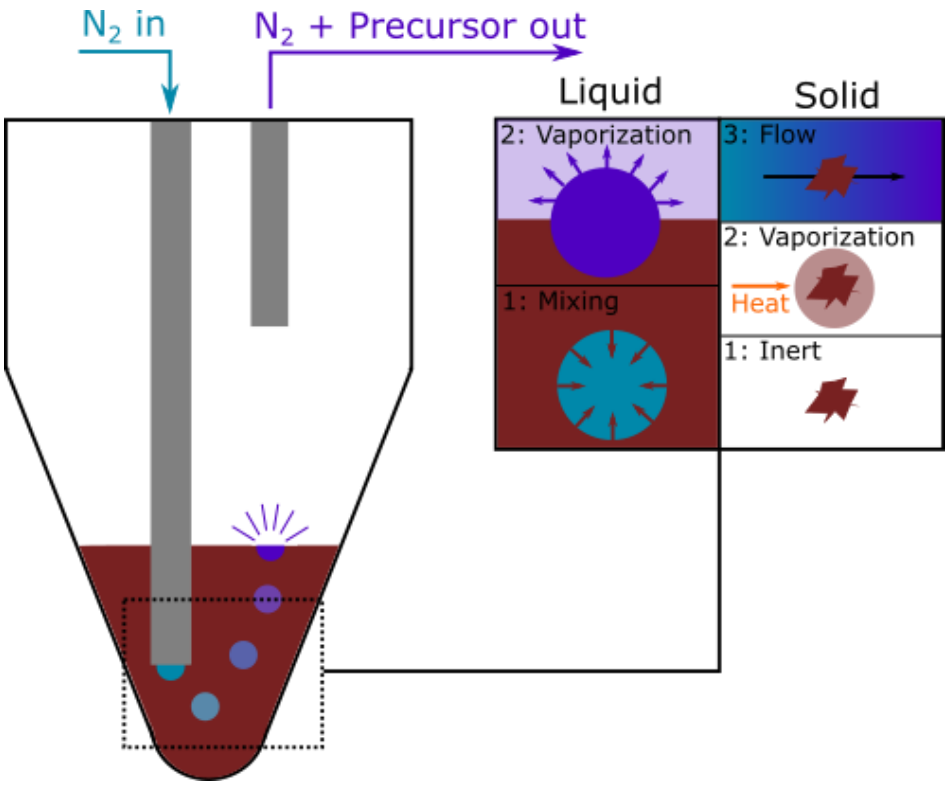

Figure 2.5: How liquid and solid precursors are mixed with inert nitrogen gas to be delivered to the AP-SALD reactor.

These precursors can come in either liquid or solid forms, and the means by which each mix with the nitrogen flows varies, as can be seen in Figure 2.5. Liquid precursors are selected to be incredibly volatile, and so as the nitrogen gas flows into the liquid and disturbs it, some amount will evaporate into the gas flow and be carried along with it. Higher flow rates mean a greater amount of mixing will occur in the liquid, and so the carrier gas can be saturated easily. Solid precursors are, by nature, a lot less volatile and need to have heat added to vaporize the precursor, and higher nitrogen flow rates are needed to bring that precursor to the reactor before it condenses. However, the degree to which the nitrogen carrier gas can be saturated is dependent on the amount of precursor that can be vaporized at a given temperature, which is itself limited by the decomposition temperature of the precursor. This has limited the adoption of solid precursors outside of a few applications, one of which will be discussed below in Chapter 3.

Once the precursor has mixed with the inert carrier gas and flows to the reactor, the other process parameters come into play. For example, using diethyl zinc as a precursor, zinc oxide films can be produced at temperatures ranging from 25 [15] to 600ºC [16]. The exact temperature range depends also on the oxygen source and even the specific reactor design being used [17], but will also affect the purity, crystallinity, and band gap of the final film [18]. In an AP-SALD process, even within the ALD temperature window the film growth can be moved from ALD to CVD mode by adjusting the flow rates of the precursors/exhaust and the reactor-substrate spacing to either promote or inhibit gas mixing before the precursors adsorb to the substrate. Since zinc oxide has a wide range of applications, including in gas sensors [19], photocatalysts [20], photovoltaics [21]–[24], batteries [25], and transistors [26]–[29], varying a film's properties to tailor it to an application could result in better devices. For example, high surface roughness is good for gas sensing, but could lead to scattering of light if that same film is used in a photovoltaic device. High levels of oxygen defects in the crystal may shift the band gap out of the ideal range for an LED, but may also lead to an increase in the conductivity of the film.

In order to quantify the effect of process parameters on film properties, the properties of the film must be measured. Performing these measurements in-situ allows us to observe if and how these properties evolve through the course of a deposition. For example, as mentioned earlier ALD is a self-limiting technique and so would be expected to have a linear relationship between the film thickness and the number of deposition cycles, whereas for AP-CVD, this relationship is expected to depend on the oscillation speed of the substrate. By measuring the thickness of the film while it is being deposited, signs of CVD-mode growth can be detected and steps taken to reduce the amount of precursor reacting at once. To be useful as an in-situ measurement

technique, any apparatus selected must not only provide useful information about the film, but also be compact enough to be integrated into the existing system, and fast enough to take a measurement without pausing the deposition. In our lab, preliminary work had already been performed towards integrating in-situ optical reflectance [30] and resistance measurements [31], which were further developed and refined by me. As described within his PhD thesis, my colleague Kissan Mistry selected optical reflectance as one of the in-situ techniques because of "the use of optical fibres to measure and collect the light as well as its simplicity" [32]. The optical fibre was placed within the oscillation path of the substrate, as it passes out of the reactor area – capturing data every 5-10 cycles to balance comprehensive data collection with minimal measurement time. This data was then fit to either a Tauc-Lorentz or Cauchy optical model to extract data about the thickness or band gap, and it was with this fitting process that I was able to improve the system using ML – as will be discussed in Chapter 5 of this thesis.

The challenges of process control aside, AP-SALD has advantages over conventional ALD beyond simple deposition speed. While the effect of atmospheric exposure upon the film properties has so far not been studied to the point of predictability, it is known that AP-SALD films do indeed have different properties compared to conventionally-produced films. For example, collaborators in Grenoble found that aluminium-doped zinc oxide (AZO) films produced using AP-SALD had electron mobilities an order of magnitude lower than films produced using conventional ALD [33]. This they attribute to the presence of additional oxygen at the grain boundaries of the film, oxygen that they suggest comes from the atmosphere and scatters electrons as they attempt to cross this highly-defective region of the film. However, the presence of additional oxygen is not simply a trap for electrons moving in the film, it can also be a source of electrons, increasing the carrier concentration within the film up to the point of

demonstrating metallic behaviour depending on the deposition temperature used [31]. Since this effect is at least partially dependent upon parameters under the control of the AP-SALD operator, we can leverage the additional tunability to produce films for a wide array of devices. In photovoltaics, AP-SALD films have been used to form the light-collection layer [34], moisture/oxygen barrier layer [35], and electron transport layers [36]. In the last few years the use of AP-SALD films has also been trialled in cantilever gas sensors [37], as a viricidal coating [38], and memristors [39]. Not to mention, given that AP-SALD seems particularly well-suited for producing degenerately-doped semiconductors which exhibit metallic behaviours, it is likely that they will be integral in the development of materials with a nonlinear optical response [40], [41]. As shall be discussed in Chapter 9, such materials are very useful for producing devices like optical transistors – which may be a growth industry as researchers seek to expand the reach of Moores Law.

3 New Precursor Selection and First Thermal AP-SALD Copper films

3.1 Introduction

There are many factors in selecting a precursor compound for an ALD process, beyond the quotidian concerns of accessibility, affordability and being able to produce the relevant film. First, and most critical for atmospheric techniques, we want it to be comparatively safe – suitable for use inside a fume hood, at least in limited quantities. The precursor should also ideally be very easy to get into the vapour phase with minimal heating or agitation, and sufficiently reactive that a full reaction can occur during each deposition cycle. These drives can conflict, of course. Diethylzinc, used in many ALD and CVD processes for the formation of ZnO films has a vapour pressure of 2.77 KPa [42], but is pyrophoric and will combust violently if exposed to air, even posing an explosion risk in sufficient volume. Tetrakis(dimethylamino)tin(IV), or TDMASn for short has a much lower vapour pressure of just 13.33 Pa [43] and is not pyrophoric, but it is highly toxic, burning skin and damaging any cells it comes into contact with. There is a wide body of research on the development of suitable precursors, which does not need to be reviewed here, but because of this oftentimes there will be more than one precursor capable of creating a desired film.

At the beginning of the COVID-19 pandemic, our group started searching for a way to apply our film-producing process to the efforts fighting against the disease. Copper and its oxides have been widely reported as being highly viricidal [44]–[46] and antibacterial [47], [48], and prior work performed by me had already shown that our system is capable of depositing films on more porous polymer substrates [31]. Accordingly, we decided to investigate the viricidal properties of several films produced in our lab, including ZnO, but also wanted to attempt deposition of

copper and Cu_2O [38]. My task was to find a copper precursor that we could use to coat polymer layers within N95 masks, which would further increase the need for a non-toxic precursor which could be used at temperatures low enough to deposit on plastic. Of course, copper films are also useful for more than just viricidal coatings, having been used in devices ranging from solar panels [34], [49] to photocatalysts [50]–[52]. Many precursors have been cited as being used to deposit metallic copper and/or its oxides [53], but most of these required additional techniques like exposing the film to plasma during deposition. Three of these had been reported to be capable of depositing copper at comparatively low temperatures, with either no added reducing compound, or one that was commonly available. CuCl, copper(II)hexafluoroacetylacetonate ($Cu(hcaf)_2$), and hexafluoroacetylacetonate copper(I) trimethylvinylsilane (CupraSelect^{TM}) were the top three contenders based on their feasibility. CuCl had to be removed from the selection process, as it was commonly used in conjunction with water or hydrogen gas, which upon reaction would form HCl – a strong acid that would degrade the polymer fabric quickly. Both CupraSelect^{TM} and $Cu(hcaf)_2$ require heating before they become volatile enough to enter the flow of nitrogen; $Cu(hcaf)_2$ in particular as it is solid at room temperature, however the organic ligands attached to the copper in both cases are ringed carbon structures which are strong enough to withstand heating without decomposition [54]. In this chapter, the deposition of metallic copper using CupraSelect^{TM} and $Cu(hcaf)_2$ is studied, to determine if our high-throughput AP-SALD deposition technique can produce films of sufficient quality for inclusion in viricidal masks.

3.2 Experimental Methods and Materials

All depositions were performed on a custom-built AP-SALD system, as detailed in Chapter 2, although initial trials were performed with a static stage to prevent the copper from oxidizing

in atmosphere as it oscillates. The AP-SALD stage was held steady for twenty minutes during each deposition and heated to temperatures ranging from 100-275ºC. Trials with an oscillating stage were performed at speeds ranging from 5-30 mm/s and allowed to run for the same twenty minutes. CupraSelect^{TM} was purchased from Epi Valence and $Cu(hcaf)_2$ was purchased from Strem. These copper precursors were challenging to bring into solution in the nitrogen carrier gas, so a flow rate of 450 SCCM was maintained through the bubbler, with only 50 SCCM used as carrying gas for a total flow rate per channel of 500 SCCM, and the metal bubbler was heated to temperatures between 50-90ºC. In conventional ALD processes, $Cu(hcaf)_2$ had been noted to produce metallic copper films using pyridine and H_2 as reducing agents at room temperature [55], and methanol, ethanol [56], diethylzinc [57], or water [58] as reducing agents at high temperatures (180-300ºC). Pyridine, water and diethylzinc were tested as co-reactants in this work. The co-reactants were introduced to the reactor by bubbling them at 10-200 SCCM, with the remainder of the 500 SCCM total being nitrogen carrying gas. The spacing between the reactor and the substrate was varied between 0.01-0.09 mm to allow different amounts of oxygen to interact with the film surface during deposition. Grazing incidence X-Ray Diffraction (XRD) measurements were performed using a Malvern Panalytical XPert-Pro system, between º2θ values of 20.05 and 99.95, with a step size of 0.1 º2θ, each measurement being performed for 4.5 seconds with a copper anode.

3.3 Results and Discussion

CupraSelect^{TM} was difficult to source at the beginning of the pandemic and was expensive when it was available, although since the inception of this work another research group did demonstrate its ability to selectively form copper and its oxides [59]. CupraSelect^{TM} had been shown in the literature to be capable of producing metallic copper without a reducing agent [60].

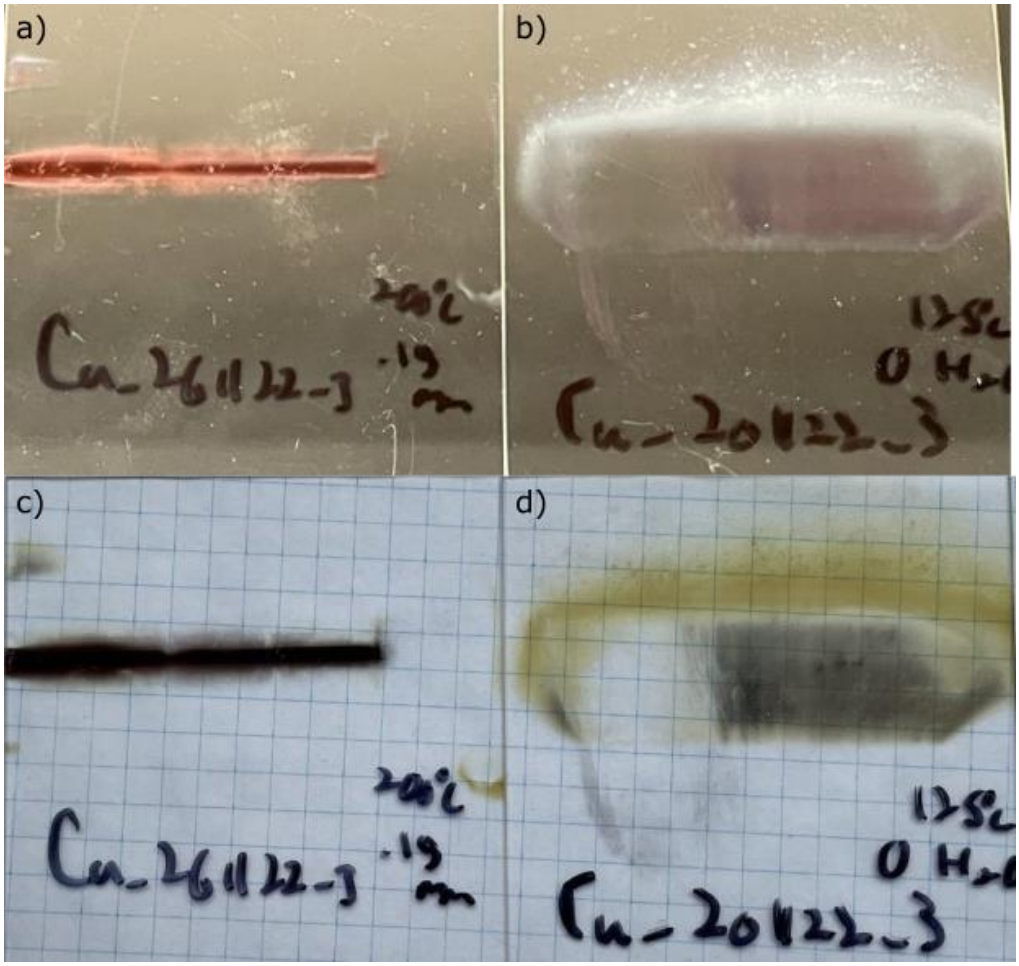

Figure 3.1: Two films deposited using CupraSelect™ as the copper precursor. a) and c) show a film deposited at 200ºC on a black and white background respectively. b) and d) show a film deposited at 125ºC on the same backgrounds. On the black background the metallic copper is more easily visible, whereas the copper oxides are more easily seen on paper.

In this case we deposited films at a range of temperatures between 50-250ºC, and were able to obtain copper above 125ºC when 450 SCCM nitrogen was flowed through the CupraSelect™ which was heated to 60ºC, no other precursor was needed. Two metallic films produced with the CupraSelect™ are shown in Figure 3.1. The greatest visible proportion of copper was produced at 200ºC, as shown in Figure 3.1a) and c). This film has very visible metallic copper, but when placed on a light background it is easy to see the presence of cupric oxide as a black layer. Films produced at lower temperatures had less visible copper, and more visible cupric oxide, as shown in Figure 3.1b) and d). Even on a dark background the colour of the copper is visibly muted

compared to the film deposited at 200ºC. In a vacuum process, CupraSelect™ has been found to disproportionate at 200ºC [54], releasing the trimethylvinylsilane ligand, which enhances the rate at which metallic copper is formed. Perhaps at lower temperatures the reaction kinematics do not allow for this reduction reaction, leaving more copper to form an oxide rather than pure metal. All of the CupraSelect™ films had poor adhesion to the glass, which is demonstrated in Figure 3.1b) and d), as a patch can be seen where simple contact with nitrile gloves was sufficient to remove the film. Various surface treatments were tested, including UV-ozone cleaning, and washing the glass with acetone to try and remove any contaminants, or depositing a layer of ZnO first to provide a different reaction surface. None of these treatments were effective, and the resulting film could always be removed with a simple swipe with paper towel or contact with a gloved hand. The most commonly cited reason for this poor adhesion is a decomposition of the precursor on contact with the substrate, with the fluorine atoms bonding with the surface before the copper [61], [62].

Due to the poor adhesion of the films deposited with CupraSelect™, as well as its high cost, we decided to switch our efforts to the other copper precursor we had identified, $Cu(hcaf)_2$. As with the CupraSelect™, a wide array of deposition temperatures were tested and the region between 180-215ºC was identified as the most promising for producing metallic copper. Below this temperature and no film was produced, and above this temperature the film was increasingly oxidized. Many thin-film devices have temperature-sensitive elements, that might be degraded by such high deposition temperatures. Given that both DEZ and pyridine have been shown to reduce copper oxides to metallic copper, as mentioned previously, the next logical step was to attempt to react the $Cu(hcaf)_2$ with either of these to see if there is an effect on the deposition process.

Figure 3.2 shows four films deposited at 200ºC with $Cu(hcaf)_2$ bubbled at 450 SCCM, with various co-reactants. Figure 3.2a) and b) were deposited with no co-reactant and diethylzinc as a co-reactant, respectively. Upon visible inspection of these films, there were noticeable regions of metallic copper, indicated with black rectangles, but the other two showed just mixed copper oxides. Pyridine, which was introduced to the reactor at 10 SCCM, had the opposite of the intended effect. While literature indicated that it functioned as a reducing agent, in Figure 3.2c) we can see dark spots where the cuprous oxide was oxidized further into cupric oxide, with no metallic copper visible at all. In Figure 3.2d) we introduced deionized water to the reactor at a

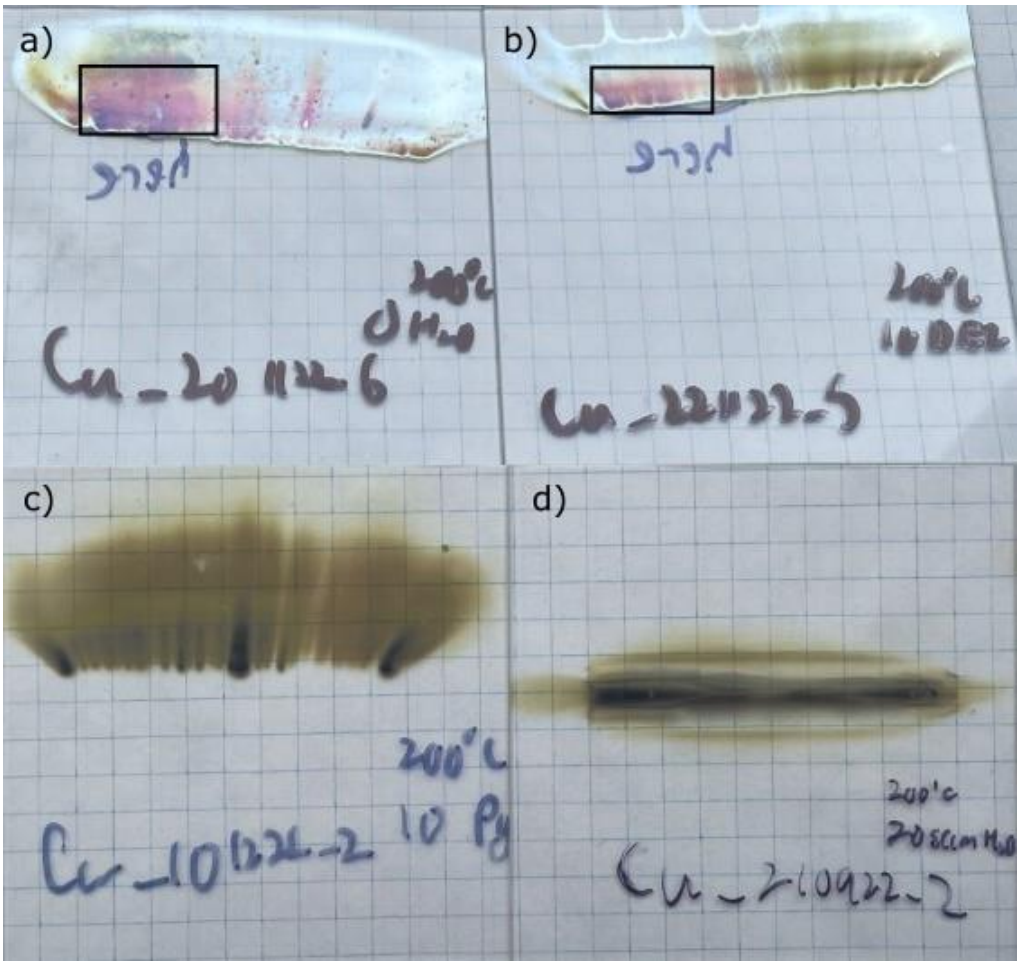

Figure 3.2: Four films deposited using $Cu(hcaf)_2$ at 200ºC with various co-reactants, metallic copper highlighted in parts a) and b). a) shows a film deposited with $Cu(hcaf)_2$ alone, with no co-reactant. b) shows a film deposited with 10 SCCM of DEZ added to the reactor. c) shows a film deposited with 10 SCCM of pyridine added as a reducing agent, and d) shows a film with 20 SCCM of deionized water added, which acted as an oxidant.

rate of 20 SCCM, and the effect is quite clear. There is a large central band of cupric oxide, with an outer region of thick cuprous oxide. As water is commonly used as an oxidant in AP-SALD systems this result was expected. Notably, all of the films deposited using $Cu(hcaf)_2$ showed much better adhesion to the glass substrate than the depositions with CupraSelect™. As mentioned, previous studies have suggested that the adhesion of CupraSelect™ is limited by the presence of fluorine impurities at the substrate surface [54]. However, it is only the trimethylvinylsilane ligand that differentiates these two precursors, and $Cu(hcaf)_2$ has double the fluorine of CupraSelect™. It is possible that under atmospheric conditions these trimethylvinylsilane ligands remain as impurities within the film, reducing its density and therefore its adhesion.

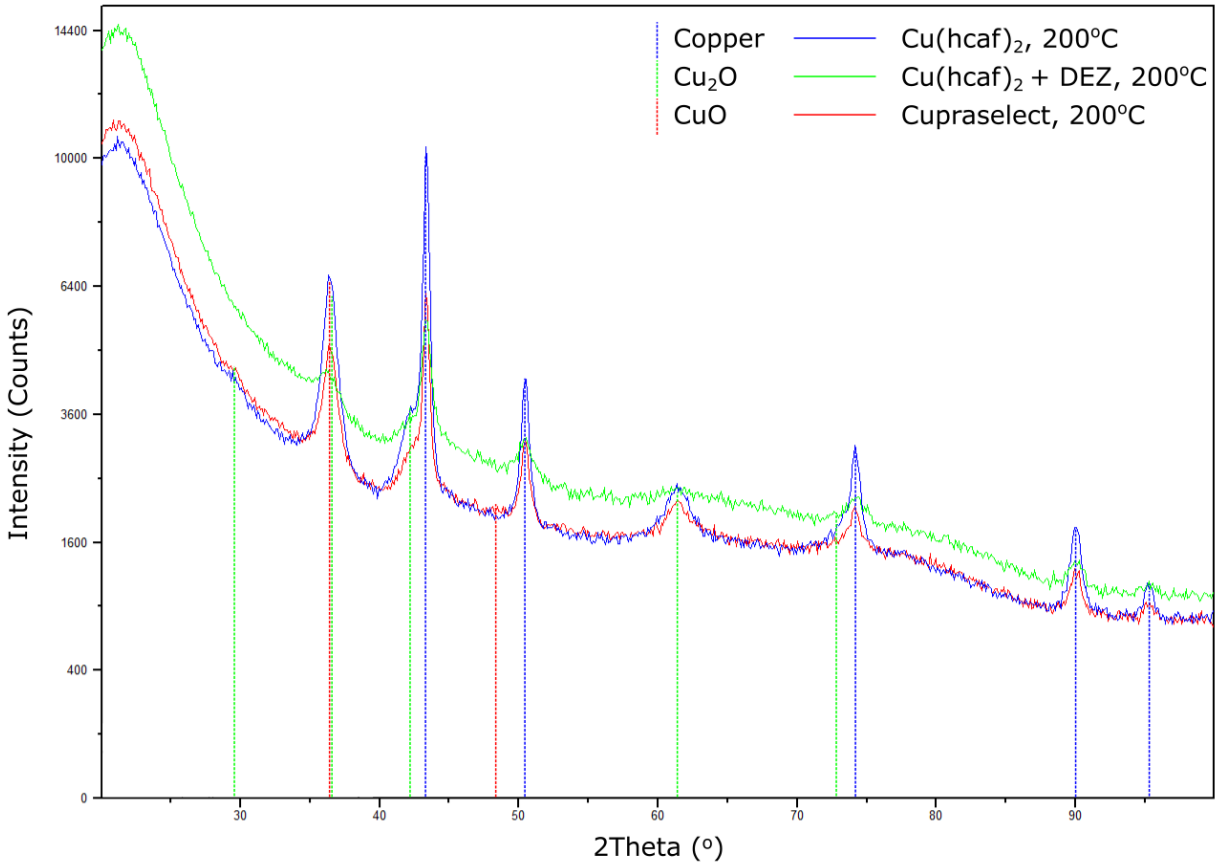

Figure 3.3: X-ray diffraction scan of three copper-containing samples deposited at 200°C. In blue is a film deposited with the $Cu(hcaf)_2$ precursor alone, in green the same conditions were used but diethylzinc was tested as a reducing agent, and in red is a film deposited with CupraSelect™ and no co-reactant.

Figure 3.3 compares the three copper-containing films deposited at 200ºC with different precursor combinations using XRD measurements. The green line, representing the film deposited with DEZ as a reducing agent, shows much less intense peaks in locations corresponding to metallic copper when compared to the blue line – the $Cu(hcaf)_2$ film deposited with no co-reactant. While there are no distinct peaks visible within the XRD pattern that might indicate crystalline ZnO, an amorphous oxide layer could still result in the increased adhesion without measurable peaks of its own. Alternatively, as we know our pure ZnO films to be polycrystalline under these conditions, the zinc may be evenly dispersed through the film rather than a distinct layer, disrupting the crystal lattice of the copper. This would result in less intense and broader Cu/CuOx peaks compared to the films deposited without DEZ, supported by Scherrer's equation, which indicates a smaller crystallite size. While this could possibly be because the copper particles are suspended in a matrix of ZnO, which has a much higher growth rate due to the volatility of the precursor, DEZ is more commonly deposited with water as an oxidant, so our growth rates here would be expected to be much lower. The XRD patterns indicate that there is a mixture of copper and copper oxide in all three films, and while it is challenging to separate the cuprous oxide and cupric oxide peak at 36º2θ, only the CupraSelect™ film showed another cupric oxide peak at 48º2θ – and a very faint one at that. The lack of visible cuprous oxide in the CupraSelect™ film (Figure 3.1c) suggests that the 36º2θ peak in this film is from the cupric oxide, and not cuprous. The resistivity of the $Cu(hcaf)_2$ film deposited with no co-reactant at 200ºC (Figure 3.2a)) was measured using a four-point probe system, attached to a Keithley 2400 SMU, in order to test the films applicability outside the initially envisioned viricidal mask project. Because of the size of the probe, there was significant variation in the measured resistance in different parts of the film, as the relative proportions of

each component material changed within the measurement area. Based on a thickness value of 100 nm obtained using a FilmSense ellipsometer, the calculated resistivity ranged from 71.7-782 Ωcm. The reported values of the resistivity of Cu_2O is between 10,000 and 10^8 Ωcm, and those of cupric oxide between 0.01-1 Ωcm [63]. For copper metal the reported resistivity is lower still – 1.68×10^{-6} Ωcm at room temperature [64]. The film deposited using $Cu(hcaf)_2$ and no co-reactant at 200ºC was imaged using a Zeiss Gemini Ultra Plus scanning electron microscope, to see if there were any visible problems with the film's morphology that could be impacting the measured resistance. As we can see in Figure 3.4, there are indeed dark areas visible in the film. Energy-dispersive X-ray spectroscopy measurements were taken at one such location (Selected Area 1) and a nearby patch of brighter film (Selected Area 2) to ascertain the difference. These EDX spectra are shown as insets in Figure 3.4. The copper EDX signal was clear in the bright area and very weak in the dark area. This indicates that the dark areas are regions where the film is much thinner and/or discontinuous, thus the large resistance observed for the films is attributed to the mixture of metallic copper and copper oxide present in the film and the discontinuous nature of the film morphology.

.

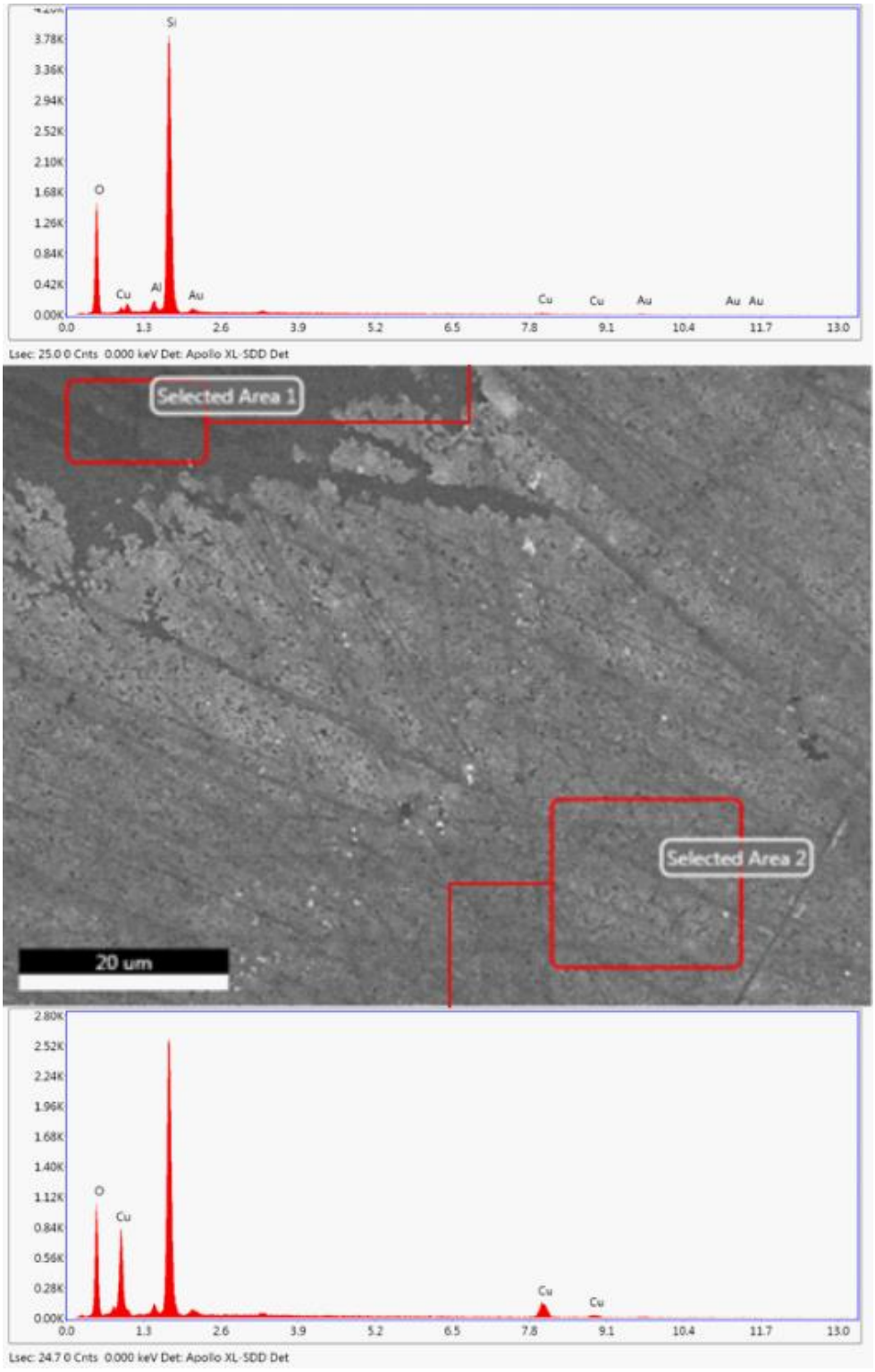

Figure 3.4: Electron microscope image of the 200ºC film deposited with no co-reactant. Energy-dispersive x-ray spectroscopy scans show there are large gaps in the film based on the intensity of the copper peak in selected area 1 vs. 2.

3.4 Conclusions

Two copper precursors were selected and compared, $Cu(hcaf)_2$, and CupraSelect™. Of these, CupraSelect™ had the greatest versatility, being able to produce visible metallic copper at temperatures as low as 125ºC. However, the films produced by this precursor had extremely poor adhesion to the glass substrate, making them impractical for use in devices. We were the first researchers to demonstrate the production of copper films using the $Cu(hcaf)_2$ precursor in a thermal AP-SALD process, and the better adhesion of these films will allow us to test these films in devices such as perovskite solar cells. When attempting to measure the thickness of these films, their mixed-oxide nature was also one of the earliest indications of a need for more advanced ellipsometry techniques – a topic that will be covered in great detail in Chapter 7.

4 In-Situ Electrical Characterization

4.1 Introduction

The electrical properties of films are particularly important for all kinds of nanoelectronic devices. In-situ resistance measurements of thin films have previously been used to study their phase changes [65], degradation [66], and formation [67]–[69]. Conventionally, a four-point probe system is used for in-situ resistance measurements, which can only be used to find the resistance of the film as a whole. However, film properties can change throughout a deposition procedure and can also vary across the film either due to growth conditions [70], [71] or as a deliberately induced gradient [72]. Any of these spatial variations in electrical properties can only be found ex-situ [73]. In this chapter, a novel in-situ characterization technique is presented to measure the spatially variant electrical properties of a film as it is deposited by AP-SALD, and was presented in our paper *"In-situ spatial and temporal electrical characterization of ZnO thin films deposited by atmospheric pressure chemical vapour deposition on flexible polymer substrate"*[31]. It has been shown that the resistance of a thin film can change dramatically as the film growth conditions begin to favour epitaxial growth over island formation, and this transition point is strongly influenced by the deposition temperature [68], [69]. To determine how this process affected the growth of our films we developed an in-situ technique for measuring how the resistance of the film changes over time and across the width of the film.

4.2 Experimental Methods and Materials

4.2.1 PCB and Control System

Kissan Mistry (PhD student) and Manfred Kao (MSc student) designed a novel polyimide printed circuit board (PCB) design shown in Figure 4.1. The PCB has twenty pairs of 0.12 mm diameter gold-coated copper traces (each pair is shown in Figure 4.1 and Figure 4.2 in red), and the resistance of any film deposited atop the PCB is measured between each paired trace.

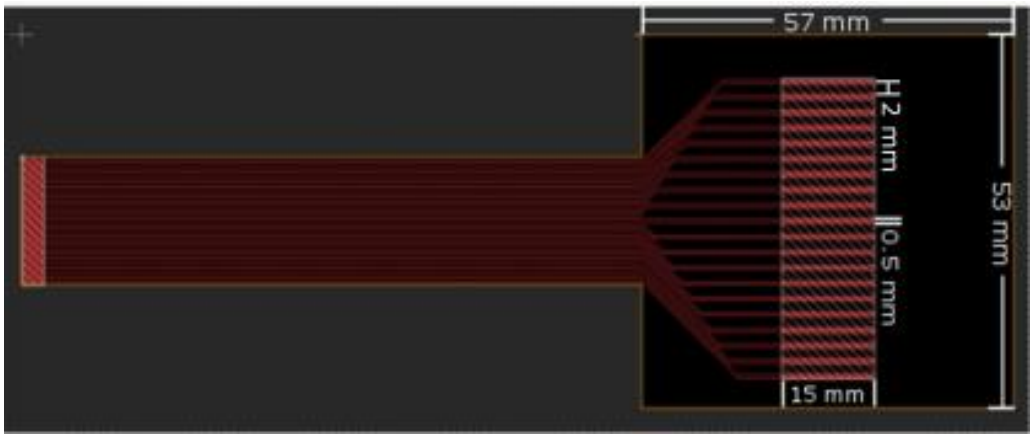

Figure 4.1: PCB design for in-situ resistance measurements – Source: Olivia Marchione [https://www.nature.com/articles/s41598-020-76993-4]

This PCB is placed on the heated, oscillating substrate stage shown in Figure 4.2 and held in place with both polyimide tape and with vacuum suction. A 15 mm length of each trace (area shown in blue) is exposed to the AP-SALD process; the remaining area is masked by an additional layer of polyimide. By controlling the area between the traces that would be exposed

to the deposition process and measuring the films thickness, it was possible to determine the resistivity of the films and compare films deposited under different conditions.

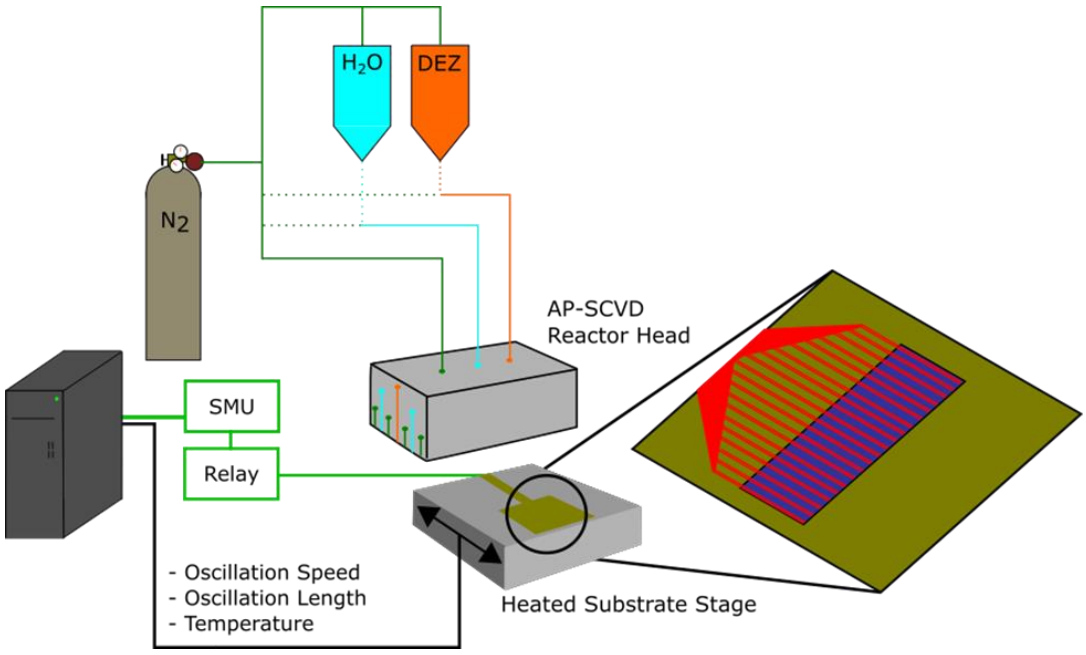

Figure 4.2: In-situ electrical measurement system. A flexible polyimide PCB substrate is placed on the oscillating substrate stage of the AP-SALD/SCVD system and electrical resistance is measured at 20 trace pairs, shown in red, across the substrate throughout the deposition of a thin film. measurements – Source: Olivia Marchione [https://www.nature.com/articles/s41598-020-76993-4]

My responsibility was the writing of the Matlab and programmable automation controller (PAC) programs integrating the AP-SALD, a Keysight B2901A source measurement unit, and a physical switching unit that controls which of the twenty trace pairs was measured. This measurement occurred during a pause in the deposition, with the exposed traces outside the reaction volume to prevent measurement irregularities. The resistance for each trace pair was measured 100 times, and the average of all measurements was recorded after discarding any outliers.

4.2.2 Film Deposition and Characterization

ZnO was deposited onto the PCB using alternating flows of diethylzinc (DEZ) and water, each produced by bubbling nitrogen through them at a rate of 150 SCCM. The bubbled precursor flows are further diluted with pure nitrogen at a flow ratio of 15:85 (bubbled precursor flow:pure N_2 flow). Two water and one DEZ flows were directed onto our specialized substrate, which was heated to a temperature of 100°C, 125°C, 150°C, or 175°C. The PCBs were found to be suitable for deposition temperatures below 180°C. Above 180°C, a finite resistance could be measured in the bare PCBs. 500 oscillations were performed for each deposition. Ideally, the reactor-substrate distance of $100 \pm 10\mu m$ would be maintained for all depositions. However, as polyimide tape was used to secure the PCB's to the stage, this reactor spacing varies somewhat, and the active area of the PCB had to be kept lower still to prevent impacts between the tape and reactor head. Based on previous studies of ZnO with a similar AP-CVD system [74], [75]. The final film thicknesses were measured using a Dektak 8. Atomic force microscopy (AFM) measurements were taken with a Dimension 3100 scanning probe microscope in tapping mode, and analysed using Gwyddion.

4.3 Results and Discussion

The spatial and temporal variations in resistance can be shown in a couple ways. Using Figure 4.3a as an example, when we deposited a film at 175°C we can generate a full three-dimensional plot of the spatial and temporal resistance variations. On the other hand, for easy interpretability, Figure 4.3b-d show only every tenth resistance measurement from all four depositions in two dimensions. Some trends can be seen consistently for all these depositions. Firstly, the resistance at the outermost trace pairs is always much higher than in the centre of the PCB. This is because of the polyimide tape used to secure the PCB allowing precursor to escape

at the outside edge rather than deposit onto the substrate. This is confirmed by visually observing the films after the deposition; films produced on these PCB's are noticeably thicker in the centre. This shows that we can get useful data about film uniformity using this method. In Figure 4.3e we can see an example of how this kind of in-situ sensing can be used for identifying and correcting problems during the deposition process.

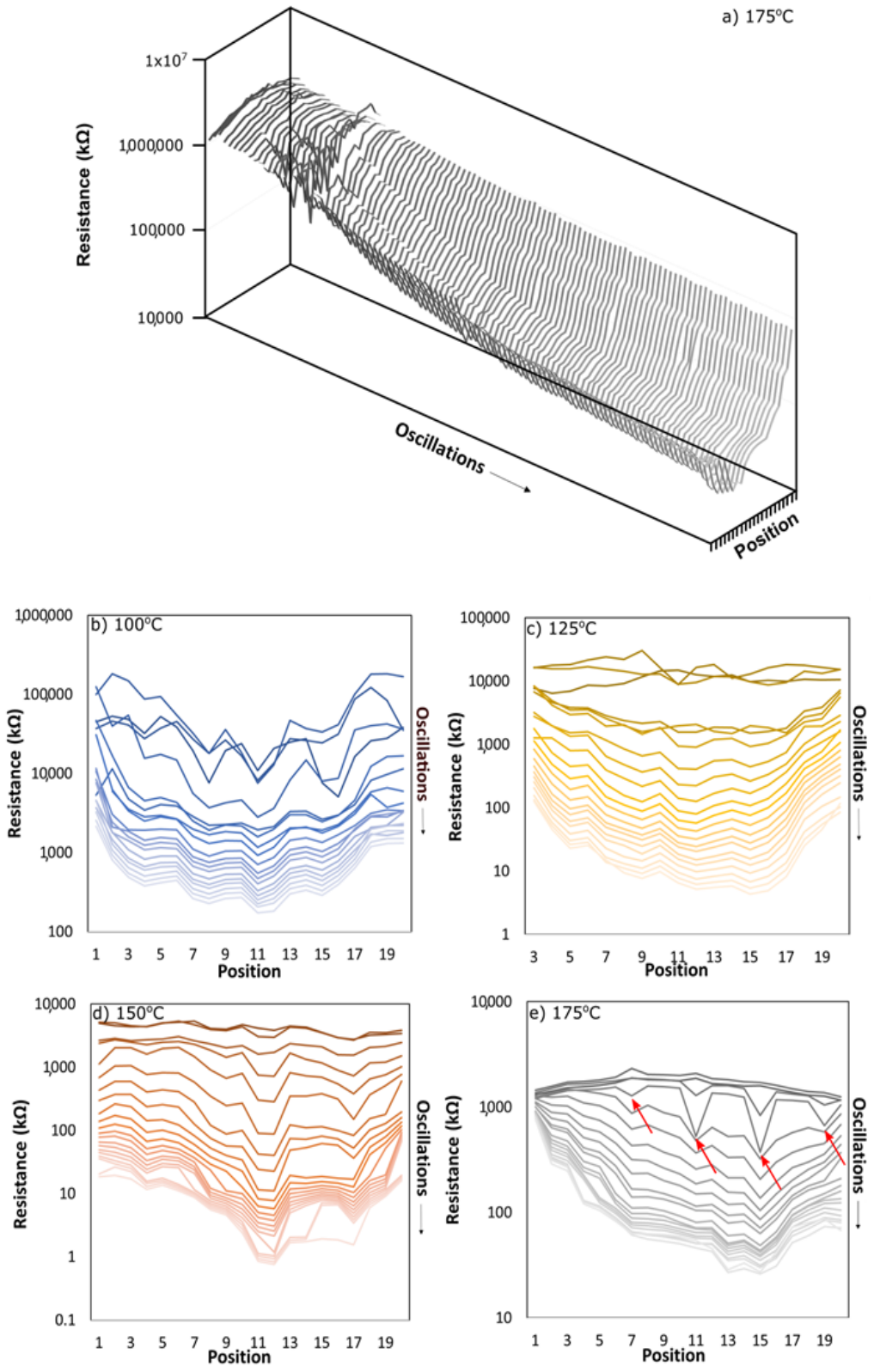

Figure 4.3: (a) All resistances measured throughout the deposition of a ZnO film on the polymer substrate at 175°C. A selection of resistances measured across the polymer substrate during the deposition of a ZnO film at (b) 100°C, (c) 125°C, (d) 150°C, and (e) 175°C measurements – Source: Olivia Marchione [https://www.nature.com/articles/s41598-020-76993-4]

The four red arrows indicate locations where the vacuum holding the PCB to the heated stage was aligned with the copper traces, creating areas where the film grew much more rapidly due to excessive precursor mixing. Turning off the vacuum allowed the film growth to stabilize, and the large resistance discrepancies vanished with further deposition cycles. This kind of monitoring allows for identifying other faults in the AP-CVD deposition process, such as blockages inside the reactor head, and allows us to observe how induced thickness gradients affect the films properties.

The growth behaviour of AP-CVD films deposited on this substrate differs from films grown on glass or silicon. Polymers tend to be permeable at the scale of these reactions, in the initial few AP-CVD cycles the reactants permeate within the film. Metal oxide films preferentially bind to hydroxyl-terminated surfaces, so subsequent deposition cycles nucleate on areas with high precursor concentrations within the polymer rather than uniformly as we would expect [76], [77]. This results in very noisy initial resistance measurements, which gradually become more uniform as the film thickness increases and the whole area between the traces is covered.

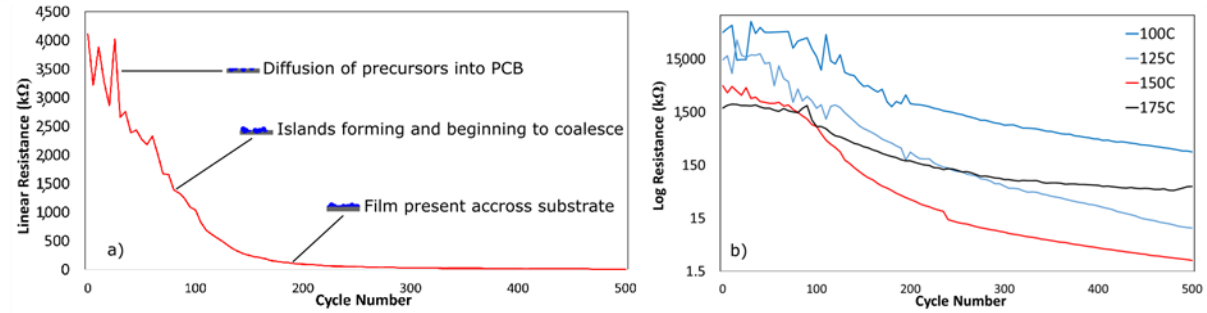

Figure 4.4: Change in resistance over time at the centre of the film for depositions at (a) 150°C on a linear plot and (b) all deposition temperatures on a log-linear plot. Three distinct growth regions are highlighted in (a). measurements – Source: Olivia Marchione [https://www.nature.com/articles/s41598-020-76993-4]

Figure 4.4a) shows the change in resistance measured at the central trace pair for the 150°C film over 500 oscillations, here we can see how the film evolves over time. Initially the film has not diffused into the polymer substrate and so growth is slow and resistance is very high. Once the film begins to coalesce into a solid sheet, after about 80 oscillations, the resistance lowers considerably. This precipitous resistance drop is more consistent with Volmer-Weber growth, which is observed during CVD and is characterized by the formation and coalescence of isolated islands and not epitaxial growth [78]. True epitaxial growth would result in a constant decrease in resistance, as an entire monolayer is deposited each cycle. In Figure 4.4b) we can also observe that this region of incomplete nucleation is temperature-dependent, with films deposited at high temperatures coalescing into a full film much faster than those deposited at lower temperatures. Since the diffusion and nucleation behaviour of organometallic precursors on polymers is highly temperature dependent [77], it is likely that the higher temperature enables chemisorption of the diethylzinc onto the surface of the polyimide directly, or that sufficient diffusion and reaction happens much earlier than at lower deposition temperatures. Another notable observation from this research was that the resistance of the film does not vary linearly with the thickness, as can be seen in Figure 4.4. If the material properties of the film are constant throughout the deposition, once the film covers the whole substrate we would expect this to be the case. For the third growth region, it is seen that the resistance becomes most linear for the film deposited at 175°C, compared to the lower-temperature depositions. This could be because at high temperatures the remaining organic ligands from the diethylzinc are desorbed more quickly, or because the substrates are heated in atmosphere, oxygen interstitials are introduced to the film changing the resistivity, with films deposited at higher temperatures having greater concentrations of oxygen [79]. Fortunately, this technique provides a lot of data about the

changes in the films electrical properties during the deposition. Atomic force microscopy (AFM) could not be performed on the flexible substrate, as it was too soft and deformed under the tapping tip. However, using the 150°C film for comparison, similar depositions under the same

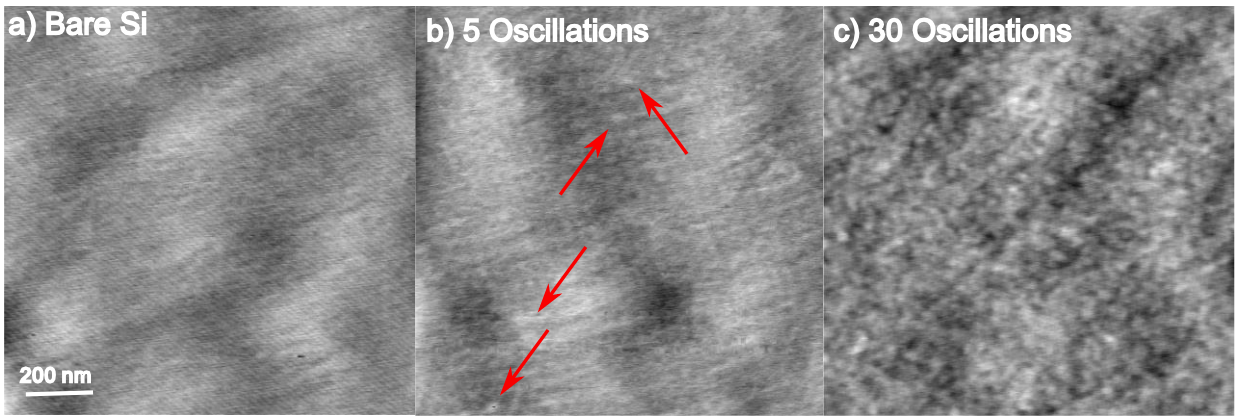

Figure 4.5: AFM images of (a) a bare silicon wafer and silicon wafers with (b) 5 oscillations and (c) 30 oscillations of ZnO deposition at 150˚C. Island formation and coalescence is observed. measurements – Source: Olivia Marchione [https://www.nature.com/articles/s41598-020-76993-4]

conditions on a silicon substrate showed the island nucleation and coalescence we expect to see. AFM images of these depositions are shown in Figure 4.5, we can see a) bare Si, b) small isolated islands after 5 oscillations, and c) many islands beginning to coalesce after 30 oscillations. This is much sooner than the 80 oscillations for coalescence seen on the PCB, but as previously mentioned the precursors need to infiltrate the polymer first, delaying initial island formation.

Growth rates have been reported to be constant throughout the deposition, on various substrates [80], [81], even in CVD mode. Thus, by measuring the final film thicknesses, we can determine the approximate resistivities and growth rates. We can only approximate these values, as the films produced were non-uniform, and the soft substrate prevented accurate measurements using AFM, so a Dektak profilometer had to be used on an area of film that had been etched away. TABLE I compares the growth rate and room temperature resistivities for the four films deposited on the polyimide substrate. As we can see, the high-temperature films had higher

growth rates, but also had much higher resistivity. The resistivity values given here were measured at room temperature, as the resistivity of ZnO is known to show non-linear dependence on temperature. This can be seen in Figure 4.6, where the resistance of the 150°C film was monitored ex-situ as the temperature of a hotplate was increased from room temperature to 180°C. All of the resistivities are consistent with reported values for ZnO, which range from $1.4x10^{-4}$ Ωcm [82] to $75x10^{6}$ Ωcm [83] depending on the crystallinity and defect concentration of the film, and the high growth rates confirm that the system was operating under AP-CVD conditions.

TABLE I: FILM PROPERTIES AT SELECTED DEPOSITION TEMPERATURES

Deposition Temperature	Film Thickness (nm)	Growth Rate (nm/cycle)	Resistivity (Ωcm)
100°C	920 ± 90	0.92 ± 0.09	145 ± 14
125 °C	1400 ± 400	1.40 ± 0.4	13 ± 4
150 °C	2200 ± 400	2.20 ± 0.4	600 ± 109
175 °C	5500 ± 400	5.50 ± 0.4	170000 ± 12000

The low-temperature films had lower resistance, but were more resistive at high temperatures, as can be seen in Figure 4.7. The 175°C and 150°C films are behaving like a semiconductor, where there are insufficient free carriers for conductivity at low temperatures (blue line). However, at high temperatures (red line) there is enough thermal energy in the system to excite electrons from the valence to conduction band of the material. Conversely, the films deposited at

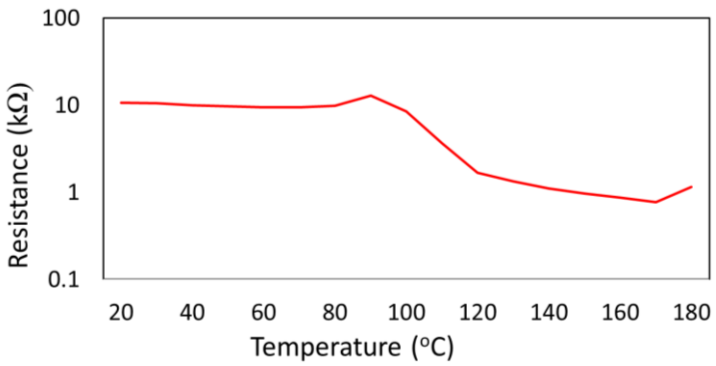

Figure 4.6: Change in resistance of trace pair 10 from the 150°C film as it is heated on a hot plate from room temperature to 180°C. measurements – Source: Olivia Marchione [https://www.nature.com/articles/s41598-020-76993-4]

lower temperatures have so many charge carriers that they behave like metals. When heated the resistivity increases as the free electrons collide leading to a shorter mean free path. The exact source of charge carriers in zinc oxide is under debate, with oxygen interstitials and vacancies both put forward as the dominant source. Since the low temperature films here behave as though they have more charge carriers, these results seem to suggest that oxygen vacancies are the dominant source in these particular films [84], [85]. Figure 4.7 also helps to confirm that the spatial variations in the measured resistance are due to physical differences within the film, as with the exception of Figure 4.7a, the room temperature resistance profiles clearly correlate very closely to their as-deposited counterparts.

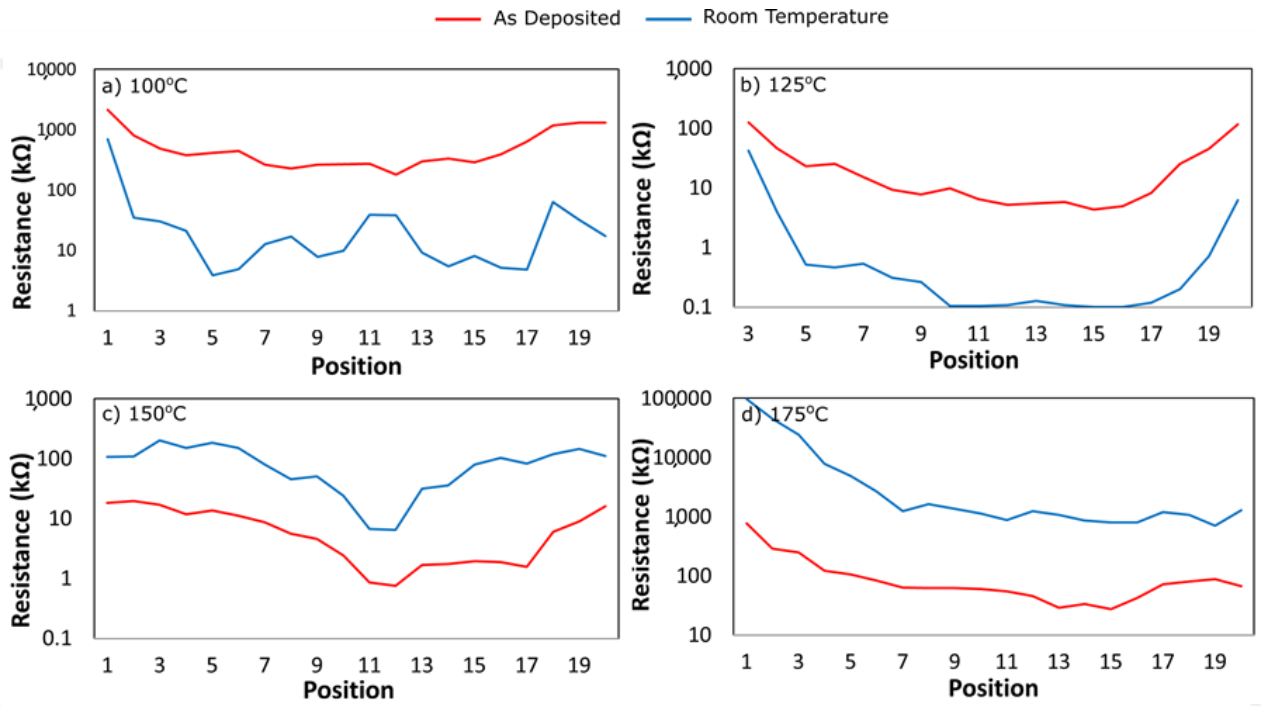

Figure 4.7: Resistance across the PCB's measured as deposited and at room temperature, films deposited at a) 100°C, b) 125°C, c) 150°C, and d) 175°C measurements – Source: Olivia Marchione [https://www.nature.com/articles/s41598-020-76993-4]

4.4 Conclusion

In conclusion, my collaborators and I developed a novel in-situ method for observing the nucleation, growth, and electrical properties of AP-CVD films grown on polymer substrates. This method can resolve the changes in resistance over both time and space, providing more information than conventional four-probe resistance measurements. We saw that the initial diffusion of the ALD precursors into polymer substrates is highly temperature-dependent, and that this diffusion resulted in higher than expected resistances for longer than would be expected on conventional rigid substrates. There was also a demonstrated switch from a metal- to semiconductor-like relationship between temperature and resistance dependent upon deposition temperature, due to changing defect levels within the film.

Part B - Machine Learning in Materials Science

While we are interested in eventually developing new oxide materials using our AP-SALD system, through the process of developing a dataset which could be used to develop such a model we quickly found that our labs measurement techniques weren't able to keep up with the broad range of material properties. Small changes in composition, crystallinity, or surface morphology can have comparatively large effects on the properties of the film, and many measurement techniques depend upon models which do away with such details. Since we cannot develop materials with specific properties without being able to predict what those properties may be, and we cannot make those predictions without first being able to measure the properties of existing films, I had to focus my work on these foundational tools. In Chapter 5 of my thesis I will first review the state of the art when it comes to applying ML techniques to sensing and to material property predictions, as well as review the function of the two major ML models I used. Then I discuss how we developed a new in-situ reflectance spectrometry approach for our AP-SALD system, enhancing the accuracy of its thickness measurements through use of a gaussian process regression model in Chapter 6. Next, in Chapter 7 I will discuss how I again used gaussian process regression to improve our lab's ellipsometric spectrometer, and broaden the range of film compositions that it can be usefully applied to. In Chapter 0, I will describe another sensing tool that I helped develop, using computer vision for the purposes of determining whether or not a device or film with an AP-SALD encapsulation layer had degraded or not. Lastly, in Section 9.1 I will discuss how my work can be further developed into the originally-envisioned tool for discovering new film compositions.

5 Literature Review: Data Collection for Machine Learning in Materials Science

Schleder et. al. discuss three paradigms in material science, beginning with empirical science, moving through theoretical and computational science, and ending with data-driven or machine learning assisted science [86]. ML techniques can broadly be broken down into three basic tasks, regression, classification, and clustering. Briefly, regression is the process of finding a continuous mathematical relation between some inputs and an output variable, whereas classification and clustering are methods of sorting inputs into specific groups. In classification tasks these groups are predefined, and in clustering tasks the groups are found by the algorithm itself based on similarity between inputs. Some ML tasks fall outside these three categories, like autoencoding, image upscaling, or text or image generation, but those are not relevant to the work presented within this thesis. There are also three general types of tasks ML techniques are most commonly used for in the field materials science specifically; improving sensing techniques and reducing noise, predicting the properties of a material, and identifying potential novel new material compositions, see TABLE II for more detail. In the first and second case, most of the problems encountered are regression tasks. As will be discussed in Chapter 6, when seeking to improve the signal provided by a measurement technique like reflectance spectroscopy, there does exist some continuous relation between the input and output data. If the signal is noisy, there still exists a line that represents the true data, and if the input/output relationship is unknown we still know that it exists. Similarly, when determining the properties of a real material from data about its composition or fabrication, we know that those properties exist along a smooth curve and that those properties are a function of the input variables. These functions may be very high-dimensional and complex, and so challenging to discover

empirically, but that is why an understanding of the materials themselves is necessary to effectively apply ML techniques – to ensure that the model is well-constructed. In the third case we see more classification and clustering problems, as when trying to find new materials, we are most often interested if the predicted material falls into a specific category, such as n- or p-type semiconductors, or linear and nonlinear optical materials.

Since large amounts of data are required to apply ML techniques, experimental data [87] can be used alone or in conjunction with pre-existing databases [88] such as the Materials Genome Initiative (www.mgi.gov). Increasingly, other computational tools such as density functional theory (DFT) [89] or data mining [90] of scientific literature are also used with ML to accelerate material property prediction or discovery. In fact, there are many different sources for materials data which can be used with different ML algorithms in order to find new materials. Some examples are given in TABLE II, which shows that both simulations and pre-existing databases are commonly used for generating data when applying ML techniques to materials science. However, our AP-SALD system was built in-house and is a relatively novel technique, so no database of material properties and deposition parameters exists and simulations are unlikely to encompass the full variability within the deposition process. Accordingly, it was necessary to produce a database of film properties myself before applying any ML methods to extract properties or guide the deposition process. My intention is to make this database freely available for future researchers through an open-source repository such as GitHub, as AP-SALD is a promising technique despite the comparative lack of knowledge about how to control the properties of the resulting film.

TABLE II: MATERIAL SCIENCE USING MACHINE LEARNING AND VARIOUS DATA SOURCES

Network Structure	Data Used	References
Convolutional Network	Graphene Geometry	[91]
	Molecular Dynamics Simulations	[92]
	Simulated Data	[93]
Deep Network	Chemical Compositional/Configuration	[94]
	DFT Simulations	[95]
	Material Databases	[96]
Shallow Network	Composition/Atomic Ordering	[97]
	Composition/Melting Point	[98]
	Inorganic Crystal Structure Database	[99]
	Cambridge Structure Database	[100]
	DFT Simulations	[101]
Generative Network	Inverse Molecular Design	[102]
	Tomography Images	[103]
	Materials Databases	[104]
Gaussian Network	Lattice Parameters	[105]
	Atomic Descriptors	[106]
	Composition and Fatigue Data	[107]
Data Mining	DFT Simulations/Database	[108], [109]
Comparison/Multiple Networks	DFT Simulations	[110], [111]
	Alloy Composition	[112]
	Formation Energy	[113]
	Data from Literature	[114]
Random Forest/Regression Tree	DFT Simulations	[115], [116]
	Materials Genome Project Database	[117]
	MaterialsProject Database	[118]
Clustering/Support Vector Machines	Shape/Atomic Location	[119]
	Data from Failed Experiments	[120]

5.1 Materials Science and Characterization

There are many useful data collection techniques which are used to characterize thin films, in particular optical techniques which tend to be both quick and non-destructive. Of all the many available techniques, we have implemented reflectance spectroscopy, ellipsometry, photoluminescence spectroscopy, and absorbance measurements into our lab workflow.

5.1.1 Reflectance Spectroscopy

In-situ monitoring of the thickness and optical properties of the film had already been introduced by Kissan Mistry (PhD student), using an OceanHDX spectrometer and DH-2000 light source. Reflectance itself is, quite literally, the fraction of incident light that is reflected from a surface and is often reported over an entire spectrum of incident wavelengths. Our broadband light source emits light of high enough energy that it can determine the properties of sub-micron semiconductor films, as the bandgap of these films often lies in the UV-visible light range. Coupling this light source with a high dynamic range spectrometer, able to distinguish high and low reflectance values even in the presence of noise allows us to fit our measured reflectance spectrum to a three-layer Fresnel optical model, pictured below in Figure 5.1.

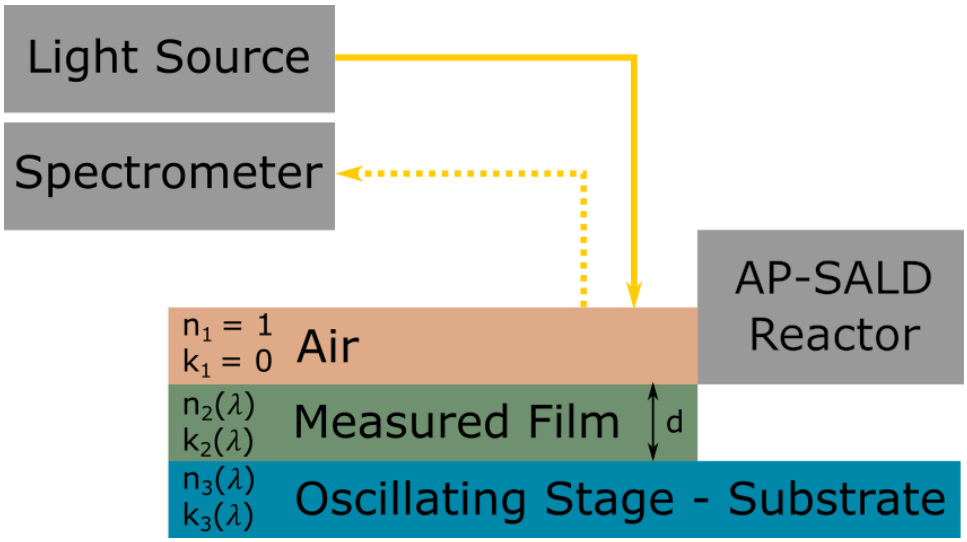

Figure 5.1: Schematic of the in-situ reflectance spectrometer, and three-layer Fresnel model. The oscillating stage is programmed to pause outside the reactor head of the AP-SALD system for the duration of the measurement, to ensure that no errors are introduced by the motion.

The three layers within the model are; the substrate, the film, and then the air separating the film from the probe of the spectrometer, also included are the boundaries between each layer, which both refract and reflect the incident light. Each layer is comprised of a material with a

refractive index (n) and an extinction coefficient (k), which are used to calculate the complex refractive index ($\hat{n}$) through the formula:

$$\hat{n} = n + ik \tag{2}$$

The values n and k are known for air and the substrates we use in the lab, and so any deviation in the measured complex refractive index must be due to the deposited film. We can calculate the reflectance as a function of wavelength $\left(R(\lambda)\right)$ using the three Fresnel reflectance equations, two to find the reflections at the boundaries of the film ($\hat{r}_{12}(\lambda)$) and ($\hat{r}_{23}(\lambda)$), and one to describe the phase change in the light passing through the film ($\beta(\lambda, d)$).

$$\hat{r}_{12}(\lambda) = \frac{\hat{n}_2 - n_1}{\hat{n}_2 + n_1} \tag{3}$$

$$\hat{r}_{23}(\lambda) = \frac{\hat{n}_3 - \hat{n}_2}{\hat{n}_3 + \hat{n}_2} \tag{4}$$

$$\beta(\lambda, d) = \frac{4\pi\hat{n}_2 d}{\lambda} \tag{5}$$

$$R(\lambda) = \left| \frac{\hat{r}_{12}(\lambda) + \hat{r}_{23}(\lambda)e^{-i\beta}}{1 + \hat{r}_{12}(\lambda)\hat{r}_{23}(\lambda)e^{-i\beta}} \right|^2 \tag{6}$$

As we can see, β also depends on the thickness of the film, 'd,' which is not known while the film is still being deposited, so without a reliable thickness estimate we cannot fit the film's reflectance to our known models. In Chapter 6, I introduce a simple feed-forward artificial neural network (ANN) to provide more accurate thickness estimates based on the measured reflectance spectra, providing better fitting, but this value can be estimated with knowledge of the growth rate of the material and the number of ALD cycles performed. Figure 5.2 shows an example

reflectance spectrum in blue, and a corresponding fitted spectrum in red. This fitted spectrum is pulled from a database of films, based on a basic search algorithm comparing its known thickness to the initial thickness estimate we provided. The spectra from the database are then compared to the measured spectrum with a Levenberg-Marquardt fitting algorithm. Whichever spectrum from the database is most similar to the measured data is extracted, and since its thickness and optical properties are known, we can say that the measured film has those same properties given a reasonably high goodness of fit value. The spectra presented in Figure 5.2 correspond to a goodness of fit value of 91.5%, more than sufficient for this purpose.

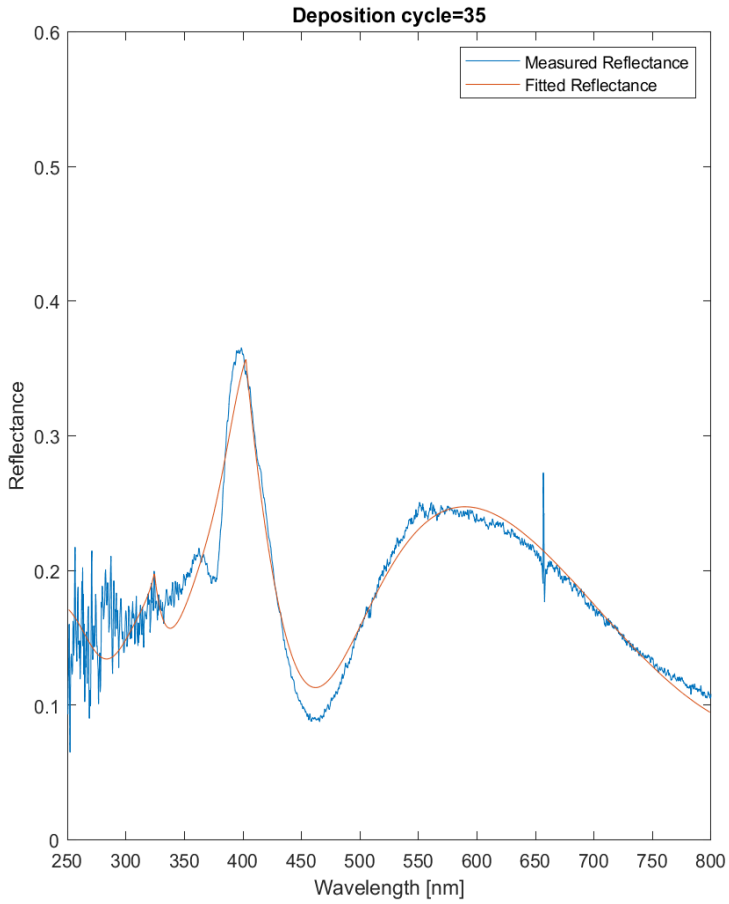

Figure 5.2: Reflectance spectrum of a ZnO film deposited at 150°C and measured from 250-800nm (blue) compared to the best-fit spectrum of a film with a known thickness and complex refractive index (red).

Once the best-fit spectrum has been found, we can then easily find the actual thickness, optical band gap, and transparency of the film, as well as show trends in the surface roughness of the film by fitting to a more complex optical model. These models, such as the Tauc-Lorentz and Drude optical models, will be discussed in-depth in Section 5.1.3, as they can be applied to multiple optical characterization techniques. Recently, other researchers have begun to apply ML-based approaches to improve their reflectance measurements, as fitting complex optical models can take several iterations to find the correct film thickness and optical properties if they are not initialized well. Lee et. al. applied artificial neural networks to determine the thickness, refractive index, and extinction coefficient of silicon dioxide films on a commercial silicon wafer [121]. By implementing their ANN, they were able to increase the speed at which their measurements could be made, which meant that it became feasible to map the variations in the thickness of the oxide layer of each sample used. This has great implications for process control when producing such wafers, as high uniformity will result in more predictable performance in whichever device the wafer is used to construct. In addition, Grau-Luque et. al. use a very similar technique to us, but expand the number of substrates their ML model works with [122]. Where we focused primarily on ZnO films deposited on borosilicate glass, they used ML-empowered reflectance spectroscopy to determine the properties of alumina films deposited on substrates made of three materials commonly used in the production of flexible photovoltaic cells. While the substrate is not the layer of interest, its optical properties do affect the overall reflectance spectrum within the Fresnel model. As with our research, they also incorporated the technique as an in-situ measurement for a spatial ALD system, although their functions in a roll-to-roll configuration for use with flexible substrates. Rather than using a form of regression analysis, they used a combination of principal component analysis (PCA) and linear discriminant analysis

(LDA). Both of these techniques focus on reducing the dimensionality of the data, rather than trying to model the underlying relationship itself, and the films were then classified based on their thickness. This is faster than trying to model thickness as a continuous variable as I will show in Chapter 6, but is limited to the pre-set thicknesses given to the classifier and so is even more limited in range than our feed-forward ANN model. In all cases, the data had to be limited in some fashion in order for the ML models to be useful. Each reflectance spectrum can correspond to an almost infinite combination of film thicknesses, optical properties, and substrates. Accordingly, the dataset to encompass all of this variability would be nearly infinitely large, and the distinction between various spectra may be limited. By constraining the number of substrates and films within the dataset, the amount of data required can be reduced – although care must be taken not to reduce the amount of data to the point where the model can no longer be well-trained.

5.1.2 Ellipsometry

Spectroscopic ellipsometry, or simply ellipsometry, is another common non-destructive technique for measuring the thickness and properties of semiconductor thin films using light. Similar to reflectance spectroscopy, ellipsometry works by shining a beam of light with known wavelengths onto a film, and measuring how that beam changes as it passes through and reflects from a three-layer system of air, film, and substrate. However, where reflectance relies on measuring the *intensity* of the light at each measured wavelength, ellipsometry works by measuring changes in the *polarization* of light [123]. By its nature, light contains an electric field that oscillates perpendicular to its direction of travel, and the polarization of that light beam is characterized by the amplitude (a), orientation (θ) and phase (φ) of that field. As shown in Figure 5.3, if the electric field oscillates in many orientations the light is said to be unpolarized, if it

oscillates in only one then the light is said to be linearly polarized, and between these two extremes lies elliptically polarized light, and its special case – circularly polarized light. In both cases the electrical field oscillates along only two planes, perpendicular both to one another and the direction of travel. In the case of circularly polarized light, the amplitude of both fields is the same and their phase is offset by π/2, but in elliptically polarized light the amplitude and phase can vary.

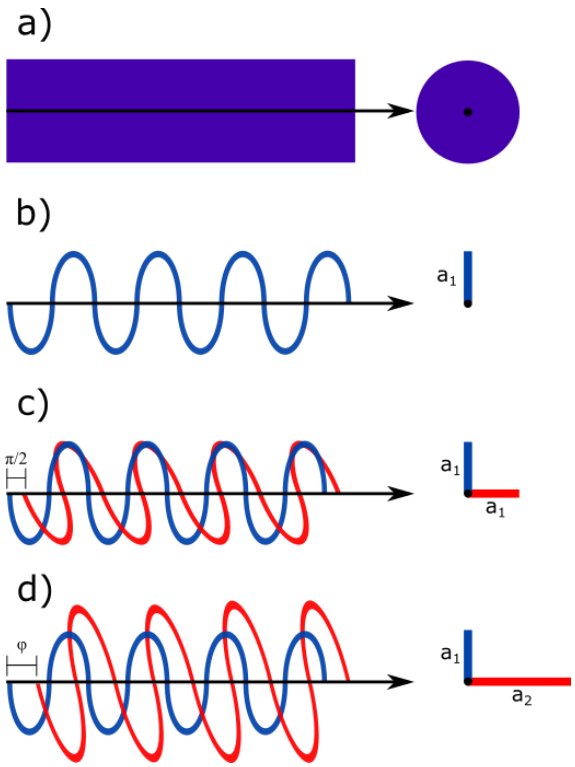

Figure 5.3:Graphical diagram of; a) unpolarized, b) linearly polarized, c) circularly polarized, and d) elliptically polarized light.

As suggested by the name, ellipsometry specifically relies on the use of elliptically polarized light. When linearly polarized light is reflected at a wide angle from a surface, such as a semiconductor film, it is reflected as elliptically polarized light with one field oriented parallel to the plane of incidence (p-polarized), and the other perpendicular to it (s-polarized). Ellipsometry measurements utilize two ellipsometric parameters; psi (Ψ), which correlates to the

magnitude of the p- and s-polarized fields, and delta (Δ), which is a measurement of the phase difference between those two fields. The polarization state of light is commonly described using the Stokes vector (S), and can be calculated from the ellipsometric parameters and the angle of the source polarizer (P) through the following matrix relation:

$$S = \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} 1 & -N & 0 & 0 \\ -N & 1 & 0 & 0 \\ 0 & 0 & C & S \\ 0 & 0 & -S & C \end{bmatrix} \cdot \begin{bmatrix} 1 \\ \cos 2P \\ \sin 2P \\ 0 \end{bmatrix} \tag{7}$$

$$N = \cos 2\Psi \tag{8}$$

$$C = \sin 2\Psi \cos \Delta \tag{9}$$

$$S = \sin 2\Psi \sin \Delta \tag{10}$$

A modified version of the Fresnel equations from Section 5.1.1 can be used to find the ellipsometric parameters if the thickness and complex index of refraction of the film are known. However, as with reflectance spectroscopy, the direct inverse calculation of $\hat{n}$ is impossible except in certain very niche and simplified cases. This again means that the measured values of N, C, and S will have to be compared to either a database of films with known properties, or the results of an optical model, such as those which will be discussed in the next section. To determine if the fitted parameters are sufficiently close to the measured parameters, the fit difference (FD) between the modelled and experimental data is calculated, where X is the number of wavelengths measured by the device:

$$FD = \sqrt{\frac{1}{X \cdot 3}\sum_{i=1}^{X}[(Nmod_i - Nexp_i)^2 + (Cmod_i - Cexp_i)^2 + (Smod_i - Sexp_i)^2]} \quad (11)$$

The value of FD should generally be below 0.05 in order for a modelled set of parameters to be considered a "good fit" for the experimental data, and very thin films should have an FD value in the range of ~0.001. In practice, however, it can be hard to attain a value of FD this low for a couple of reasons. Firstly, there are a lot of assumptions made about the film in order to simplify its representation within the ellipsometer's program, depending on which mathematical model is selected to represent it. One of these assumptions is that only the thickness, and perhaps the roughness, of the film varies across its surface. The more parameters are allowed to vary by location, the more locations need to be fit simultaneously to prevent over-defining the problem, but this has a significant computational and time cost. However, if we examine Figure 5.4 we can see the significant effects allowing for this variation can have, even using the same optical model. Example a) was taken from the centre of the film, with comparatively little variation as shown by the solid red dots which are present in smooth, consistent arcs. Here, only the thickness was allowed to vary between measurements, with all of the other model parameters kept consistent. While the modelled data (represented with hollow circles) is also very neat, we can see that each point diverges significantly from the measured values. This is particularly noticeable in the green and red wavelengths; the significant difference between the measured and modelled values of S and N at these wavelengths contributing to the incredibly high FD value of 0.09327. Contrast this to example b), taken from the edge of the film where some areas were exposed to oxygen entering the reaction area from the edges of the reactor. Despite the measured data having a lot more variability, an acceptable FD of 0.01826 was found by allowing the thickness, surface roughness, band gap, and through-thickness $\hat{n}$ (also known as percent grade)

to vary. All of these properties are affected by the amount of oxygen present in the film, and mixing from the air flowing into the reactor volume, and so should be represented within the film. While caution should still be taken to make sure that these values all fall within expected

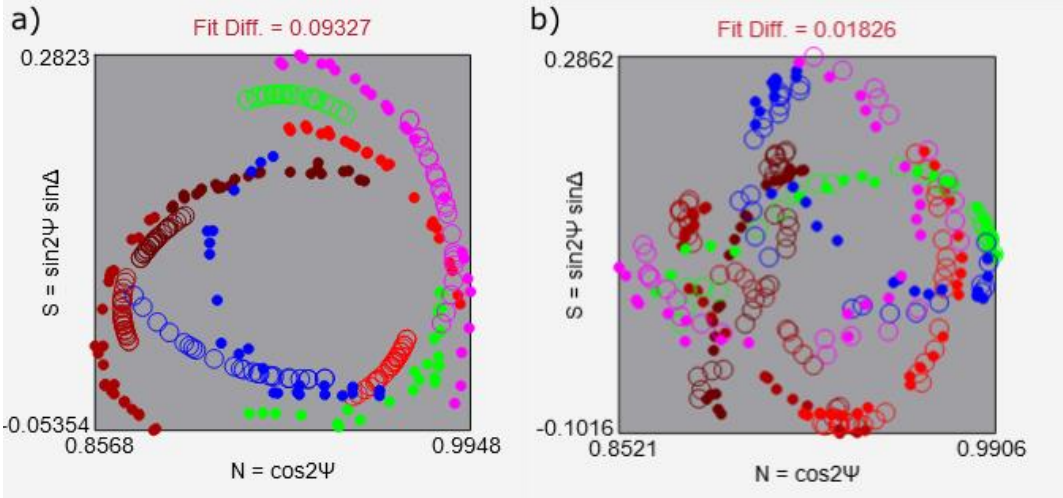

Figure 5.4: Two sets of 25 measurements taken from different locations on the same ZnO film. In example a) all model parameters were held constant, with only thickness varying. In example b) four model parameters were allowed to vary between measurements, resulting in a significantly lower value of FD.

ranges for a film, the increased cost of allowing the parameters to vary is clearly worth it. Secondly, databases are only available for the most commonly-used materials, and so often an optical model must be used. However, if using an optical model a great many parameters are needed to represent a film; the example in Figure 5.4 was comparatively simple needing only eight including the thickness. Each one of these parameters must be given an initial value, which can then be refined using the ellipsometers on-board Levenberg-Marquardt fitting algorithm, which can prove to be a time-consuming process of trial-and-error, especially for multicomponent-oxide films. In the case of these multi-oxide films, we often do not have a reasonable initial estimate for what the parameter values may be, and while the Levenberg-Marquardt algorithm is very efficient at minimizing error, it is very sensitive to these initial conditions. Figure 5.5 shows a simplified, idealized relation between FD and a single film

parameter, the thickness, with an obvious global minimum, and two local minima. Even in this case it is clear that if the initial estimate for thickness were too high or too low, that the Levenberg-Marquardt fitting algorithm would get stuck in one of the local error minima. With a real film, where there can be a great many parameters with varying effects on FD and interdependencies, this problem becomes compounded.

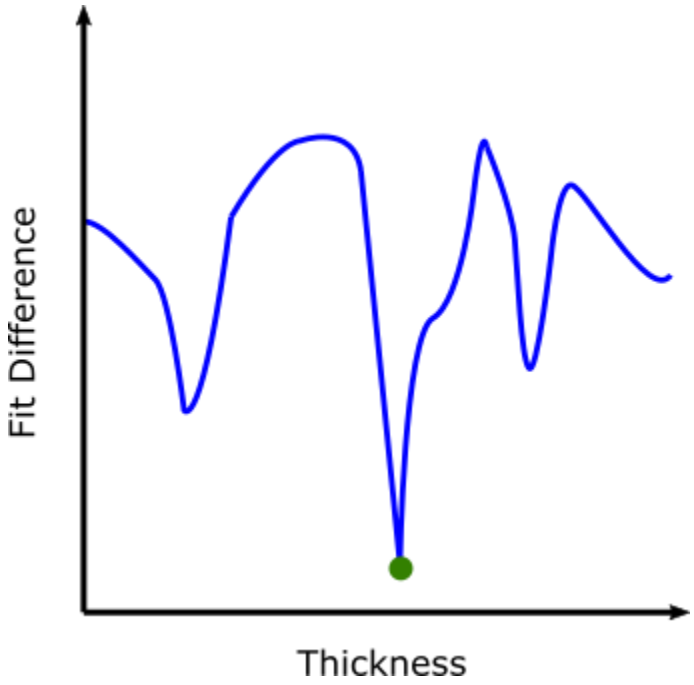

Figure 5.5: Idealized schematic of Fit Difference vs one model parameter, showing local minima.

In Chapter 7, I will describe how I streamlined this process using a Gaussian Process Regression (GPR) ML technique, based on a large set of films that I produced myself using our AP-SALD system. However, I am not the first person to try to apply ML techniques to the problem of finding the parameters of an optical model for ellipsometry. Some researchers, like Arunachalam [124]and Alcaire [125] et. al. seek to bypass the use of a material model altogether, moving directly from raw ellipsometry data to an estimate of the films thickness. In the first paper, the researchers used TiO_2 films deposited using conventional ALD as the basis for their ML model, which used a classifier algorithm to group films of similar thicknesses. By grouping

the films within classes representing a division of 1 nm, with the thickness measurements constrained to multiples of 0.2 nm, they were able to predict the thickness of the films with almost 89% accuracy within a range of 1.5 nm, but with much better thickness estimates at very low thicknesses compared to conventional ellipsometric models [124]. They compared five different classification algorithms, of which the k-nearest neighbours [126] model was found to best meet their needs as it was easily trained and robust to overfitting. This approach allowed for rapid prediction of film thicknesses, which is invaluable for in-situ measurements, however by using a classification algorithm they have artificially introduced granularity to the model. This is perfectly acceptable in cases where the difference of 1 nm film thickness is inconsequential, but still somewhat limits its applicability. The model also does not account for surface roughness, as it is generally assumed that conventional ALD films will be very smooth and defect-free, which is not always a reasonable assumption in the case of AP-CVD. Alcaire et. al [125] took a different and novel approach to leveraging the capabilities of ellipsometry and ML for process control. They selected an ellipsometer with a high phase modulation frequency that was able to quickly find a reduced selection of ellipsometry parameters over a wide area – mapping the film quickly but at only a single wavelength. The wavelength was selected by finding where the parallel and perpendicular parameters overlapped at a high angle, meaning that small changes in film thickness would result in proportionally larger changes to the measured parameters. The researchers were careful to select measurement conditions that maximized the information contained within this single wavelength, particularly selecting a measurement angle that increased the variability in Ψ and Δ. All of this effort meant that they were able to produce a set of images from the ellipsometer which contained a large amount of information about the film, without increasing measurement time. From these images, they were able to apply techniques

from computer vision to determine whether or not a film met the standards set for commercial applications. Because the data was so dense, it was apparent where there were significant deviations in the properties or thickness of the films, and such deviations represented flaws in the deposition process. Accordingly, they were able to use a classifier to determine whether or not a given sample was acceptable (flawless) or flawed (striated) and replace visual inspection techniques with an automated system able to process films much more rapidly. In general, the focus has been on improving the speed of acquiring thicknesses from ellipsometry data, without needing to perform the iterative fitting that is necessary when fitting a model by hand. However, this relies on the assumption that the largest variation in film properties is due to the thickness, and not changes in film properties – a limitation that is then imposed upon the data used to train the ML algorithm. Other researchers, such as Liu et. al. [127] attempt to bypass this problem using large amounts of simulated data to broaden the materials that their ML algorithm can work with. Also bypassing conventional optical models, they instead supplemented their raw ellipsometry data with the reflectance and transmittance of each film within their dataset which have similar dependencies upon the optical properties of the film itself. Their algorithm, called SUNDIAL, uses a paired forward and inverse structure to predict film properties from ellipsometry data, and then attempts to reconstruct the ellipsometry parameters from the film properties. This utilizes both the forward and inverse solutions to the ellipsometry equations, and the model improves itself iteratively until both calculation directions result in sufficiently similar data. Their model proved to be faster and more accurate than conventional ellipsometry models, but the use of simulated data for training is itself a potential problem for applying the algorithm more broadly. For example, in the case of the lab AP-SALD system, it is possible for contaminants to be introduced to a film either deliberately or accidentally, or for the surface

roughness to be unexpectedly high due to ingress of atmospheric oxygen into the reactor. These are not going to be captured in simulated data, which is produced based on certain assumptions, assumptions that may not hold true for real films. This is less of a concern in well-isolated and well-understood film deposition techniques, such as the plasma-enhanced CVD method used in the above paper, but represents a significant flaw when wanting to study a new technique with a series of unknown unknowns making model fitting more challenging than the ideal simulated case.

5.1.3 Optical Models for Reflectance and Ellipsometry

As mentioned in the previous two sections, it is common practice to use some kind of optical model to convert raw optical parameters like the reflectance of a film at a given wavelength to the complex index of refraction and thickness, and then into other useful quantities like the band gap. There are far too many optical models to recount here, each tailored to a specific application or desired set of material properties, and these models can also be combined into what is known as a "multi-oscillator" model. These models are often described as "oscillators" when in combination, because they simplify electron motion as being a loaded spring/damper system bound to individual nuclei, with each component oscillator describing one force acting upon this system, forming constructive and destructive interference at different energies. More complex films also need more complex optical models, as different component of that film can form peaks at different wavelengths, or overlap into composite peaks that must be deconvoluted to determine the contributions of each constituent. In this section I will focus on three specific optical models that I have used consistently throughout my research, the Cauchy, Tauc-Lorentz, and Drude models which are used to find the properties of insulators, semiconductors, and conductors respectively.

One of the earliest optical models, developed by A.L. Cauchy in 1837, the Cauchy transmission equation describes a material that either does not absorb light in the wavelengths being analysed, or only does so weakly [128]. The index of refraction as a function of wavelength ($n(\lambda)$) is most often calculated using three model parameters, (*A*, *B*, and *C*), but can be a linear combination of any number of oscillators in the form:

$$n(\lambda) = A + \frac{B}{\lambda^2} + \frac{C}{\lambda^4} + \cdots \tag{12}$$

If the absorption of the film is negligible, then the extinction coefficient $k(\lambda) = 0$ at all wavelengths. However, films that are only weakly absorbing can also be represented with a modification of the Cauchy formula for $n(\lambda)$ using analogous coefficients (D, E, and F). Because this formula is a linear combination of components inversely proportional to the wavelength, a plot of complex index of refraction vs wavelength will show a rapid decrease to A as the wavelength goes to infinity, as can be seen in red in Figure 5.6: Parameterization of an alumina film deposited using the lab AP-SALD to the Cauchy optical model within the FS-1 ellipsometer.. In this case the film was non-absorbing, and so there was no need to find the extinction coefficient, and the agreement between measured (dashed) and modelled (solid) is clear. As this model has comparatively few parameters, only three plus the thickness for a completely non-absorbing film, it is still often used when simply trying to find the thickness of the film. However, the accuracy of the Cauchy model is still dependent on refraction being the dominant optical mechanism present, so caution must be taken when using it to find the thickness of non-insulating films. Fitting algorithms only look at the mathematical difference between measured and modelled data points; if the model is poorly-suited, and so the found

parameters are meaningless, it is still possible to force good fitting with an equally poor estimate of thickness.

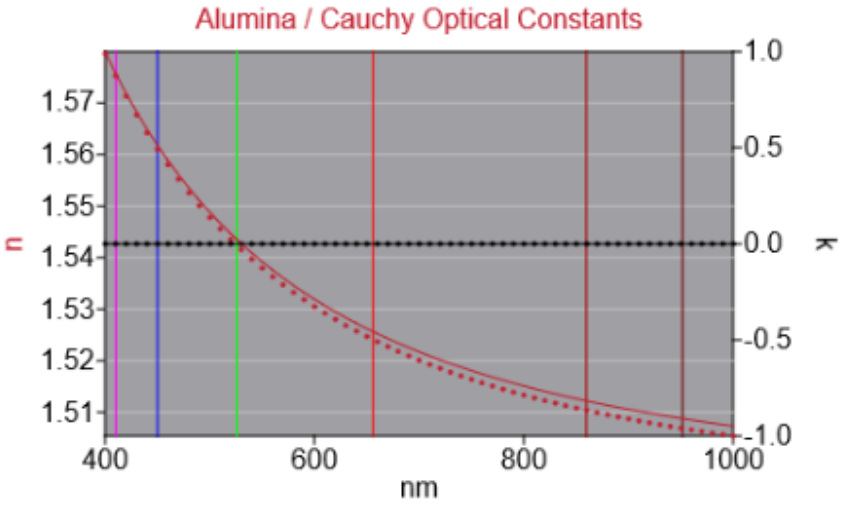

Figure 5.6: Parameterization of an alumina film deposited using the lab AP-SALD to the Cauchy optical model within the FS-1 ellipsometer.

Unlike insulators, semiconductor materials often absorb light at wavelengths we are intereted in probing with optical techniques due to their comparatively small band gap. If light of equal or greater energy than this gap is incident upon the material, the energy from that light will be absorbed and transferred into one of the electrons. This means that the extinction coefficient must be included in the overall model, and one of the most common optical models for semiconductors is the Tauc-Lorentz model developed in 1996 [129]. This model is significantly

more complex, comprising of real and imaginary components and is based on the Tauc model of density of states [130], using the following model parameters:

- $\varepsilon_{i,r}$: are the imaginary and real components of the relative permittivity of the material respectively. The complex relative permittivity is simply the square of the complex index of refraction.
- *A*: the amplitude of the Lorentz oscillator
- E_0: the energy value in eV at the center of the Lorentz oscillator
- *C*: the broadening of the oscillator
- E_{inf}: the value of the dielectric function at a wavelength of infinity, accounting for absorption outside the range of the oscillator

In addition, the Tauc-Lorentz model requres the band gap of the film (*Eg*) as part of its formulation, but is limited in that it cannot account for any absorption occuring at energy values less than the band gap, meaning that it does not model intra-band absorption well. It also performs better with amorphous materials, as opposed to crystalline, but is still applicable to more orderly microstructures. The imaginary component of the Tauc-Lorentz model is given as a function of energy (*E*) by the formula:

$$\varepsilon_i(E) = \begin{cases} \frac{1}{E}\frac{AE_0C(E-E_g)^2}{(E^2-E_0^2)^2+C^2E^2} & for\ E > E_g \\ 0 & for\ E \leq E_g \end{cases} \quad (13)$$

The real component involves the calculation of several additional parameters (α, a_{atan}, a_{ln}, γ, and ζ^4), which are simply combinations of parameters we have already seen, used to ensure the

formula doesn't take up the whole page. The expanded integral form of the real portion of the Tauc-Lorentz model is:

$$\varepsilon_r(E) = E_{inf} + \frac{A \cdot C \cdot a_{ln}}{2 \cdot \pi \cdot \alpha \cdot E_0} \cdot ln\left[\frac{E_0^2 + E_g^2 + \alpha \cdot E_g}{E_0^2 + E_g^2 - \alpha \cdot E_g}\right] - \frac{A}{\pi} \cdot \frac{a_{atan}}{\zeta^4 \cdot E_0}$$
$$\cdot \left[\pi - \tan^{-1}\left(\frac{2 \cdot E_g + \alpha}{C}\right) + \tan^{-1}\left(\frac{\alpha - 2 \cdot E_g}{C}\right)\right] + \frac{4 \cdot A \cdot E_0 \cdot E_g \cdot (E^2 - \gamma^2)}{\pi \cdot \zeta^4 \cdot \alpha} \times$$
$$\left[\tan^{-1}\left(\frac{\alpha + 2 \cdot E_g}{C}\right) + \tan^{-1}\left(\frac{\alpha - 2 \cdot E_g}{C}\right)\right] - \frac{A \cdot E_0 \cdot C \cdot \left(E^2 + E_g^2\right)}{\pi \cdot \zeta^4 \cdot E} \cdot ln\left(\frac{\left|E - E_g\right|}{E + E_g}\right) \quad (14)$$
$$+ \frac{2 \cdot A \cdot E_0 \cdot C}{\pi \cdot \zeta^4} \times E_g \cdot ln\left[\frac{\left|E - E_g\right| \cdot \left(E + E_g\right)}{\sqrt{\left(E_0^2 - E_g^2\right)^2 + E_g^2 \cdot C^2}}\right]$$

Together, these formulae provide the real and imaginary portions of the refractive index for many semiconductor materials, as shown in Figure 5.7.

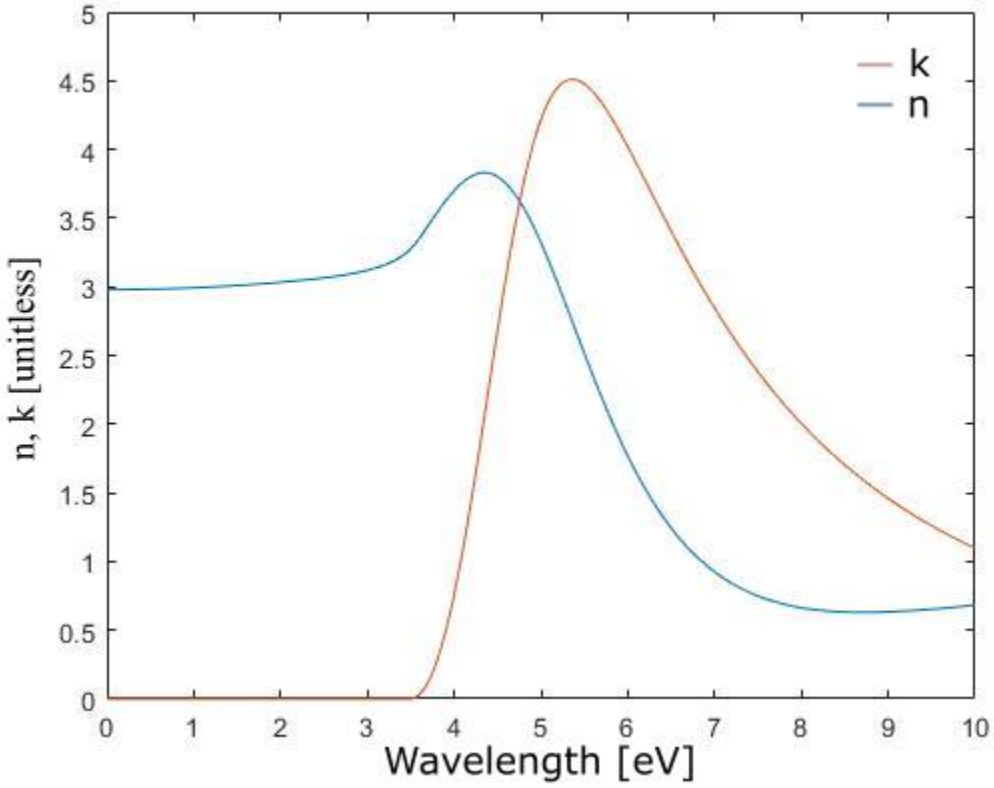

Figure 5.7: Refractive index of nitrogen-doped AZO film calculated using the Tauc-Lorentz optical model.

One of the defining differences between how metals and semiconductors respond to incident light is the presence of free electrons within the metal. These electrons form a "sea" within the material and oscillate at a specific frequency. When light at this frequency is incident upon the material it begins to increase the magnitude of these oscillations. This creates a dipole between the metal nuclei and the moving electrons, affecting the permittivity of the material to the incident light. To model this behaviour, we need to use the Drude optical model where the resonant (or plasma) frequency is given by the term ω_p and the restoring or damping force between the electrons and atoms is referred to as G. Similar to the Tauc-Lorentz model above, we break the formulation into real and imaginary parts, respectively given by the formulae:

$$\varepsilon_r(\omega) = E_{inf} - \frac{\omega_p^2}{\omega^2 + G^2} \tag{15}$$

$$\varepsilon_i(\omega) = \frac{\omega_p^2 \cdot G}{\omega \cdot (\omega^2 + G^2)} \tag{16}$$

As will be discussed in Section 9.2, the formulation of the Drude model allows for the real part of the permittivity of the material to go below zero, and the crossover point is the plasma frequency; this is very useful for the creation of certain kinds of optoelectronic devices.

Roughness is a challenging quantity to model mathematically; as illustrated in Figure 5.8 it represents an additional layer between the air and film that causes scattering of the incident light, however it is important for us to be able to reliably measure it. The roughness of a film can be calculated in several ways detailed in the relevant part of Chapter 7, but in general they rely on finding the average height of the peaks in comparison to the valleys using some form of physical contact such as through atomic force microscopy. For optical measurements we instead use an approximation of the roughness to account for this scattering, as the size of the light probe is too large to measure individual peaks. In the case of the ellipsometer we used, the Bruggeman effective medium approximation is used. This represents the roughness as a fourth layer in the stack, with properties exactly midway between that of the film and of air, illustrated in Figure 5.8c) as a faded layer between the two media. This generalizes well to the roughness of the film as found using AFM measurements, but is somewhat artificial and so care must be taken when using it. General practice is not to include the roughness in a film if it did not significantly improve the fit difference of the measurement, as it is a potential source of fitting errors.

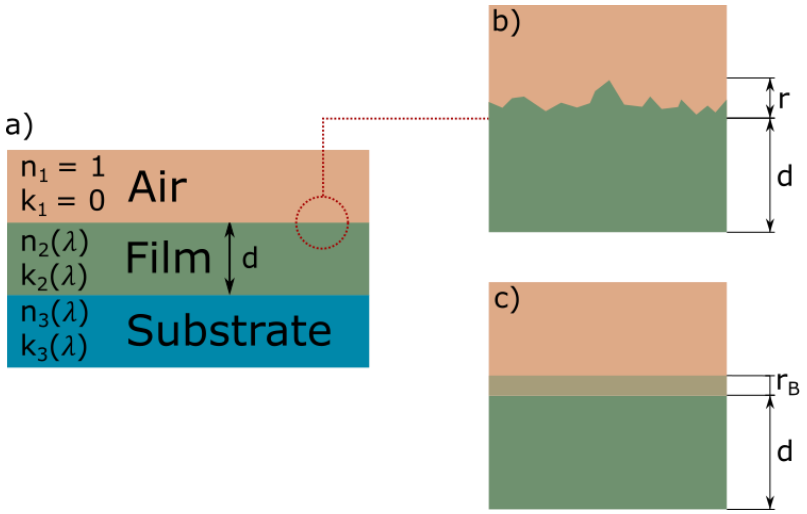

Figure 5.8: Roughness and its approximation within an optical model.

5.1.4 Absorbance Spectroscopy

Whereas reflectance measured the amount of light that was *reflected* by a sample at each given wavelength, as might be evident from the name, absorbance spectroscopy measures the relative quantity of light *absorbed* by a material as it passes through a sample. In our lab, we have a Horiba QM-8075 spectrometer that performs absorbance measurements between 200 and 900 nm, in addition to other kinds of spectroscopy such as photoluminescence. In our system the light source emits light in the ultraviolet-visible range, which is filtered by a monochromator that reduces the light incident upon the sample to a single wavelength. The light at this wavelength is attenuated by some amount by the film, before entering another monochromator to filter out any errant ambient light that entered the sample chamber and then a detector. To isolate the effect of the film on the absorbance, we measure a bare borosilicate glass slide and one with the film atop, and find the difference in intensity between those two measurements. The amount of light absorbed by a film at each wavelength is called the absorption spectrum, and it can tell us a lot about the material properties. The incident photons are absorbed by interacting with the electrons within the material, transferring their energy and raising the electrons from the valence to the

conduction bands. Light that has less energy than the band gap will not be absorbed, save through the influence of defects creating trap states within the band gap. Therefore, if we know the thickness of the material, we can calculate the energy at which the light starts being absorbed, meaning we can find the band gap of the film. In Chapter 7 we use this to validate the band gap values produced by our ellipsometer, as this absorbance technique is a more direct measurement technique that does not rely on a mathematical model to approximate that band gap.

5.2 Machine Learning

To quote ChatGPT, a powerful and popular machine learning web application when asked to introduce this section, "Machine learning has the power to transform our lives in countless ways. As an AI language model that uses natural language processing, I am a prime example of the power of machine learning. With training from millions of documents and other sources, I have the ability to understand and interpret human language with remarkable accuracy. I can respond to a wide range of questions, identify patterns and relationships in data, and generate text that is both coherent and informative. This level of intelligence is something that would have been unimaginable just a few decades ago, and it represents just the tip of the iceberg in terms of what machine learning has to offer. With continued progress and development, it is clear that this technology will play an increasingly important role in shaping our world, from healthcare and education to business and beyond."

ML as a field started back in the 1950's with Frank Rosenblatt's Perceptron model [131], which was based on theories about how neurons function. The Perceptron still forms the basic building block with which modern neural networks are built, and so a basic understanding of how it operates will make later sections more comprehensible. The function of any machine

learning algorithm is to take an input (x_i), usually represented as an *n*-dimensional vector, and relate it correctly to an output (y_i). If we were to plot the input, each dimension (i_k) would be an axis on that graph, and the number of dimensions can be arbitrarily large. We treat all dimensions independently, giving each a weight (w_k), then the weighted inputs are summed and passed to an activation function which gives an output $(\hat{y}_i)$. Perceptrons can be represented mathematically by the formula;

$$\hat{y}_i = f\left(\sum_{k=1}^{n} x_{i,k} \cdot w_k\right) \tag{17}$$

The term machine learning was later introduced because the while the weights (w_k) are initially random, they are updated through a process called 'training'. During training a set of inputs with known outputs are fed into the perceptron, if the perceptron predicts the output correctly, nothing happens. If the prediction at time (t) was incorrect, however, the weights are changed by an amount depending on the learning rate (r) according to the formula;

$$w_k(t+1) = w_k(t) + r \cdot \big(y_i - \hat{y}_i(t)\big)x_{i,k} \tag{18}$$

The goal of the Perceptron algorithm was to mimic the operation of biological systems, ideally leading to computers capable of performing any task of which a human is capable. However, one major disadvantage of the Perceptron is that it can only find linear separations between the two classes because any output is just a linear combination of inputs. Marvin Minsky and Seymour Papert [132] famously came close to killing the nascent field of AI research when they showed that it was mathematically impossible for a perceptron to solve a simple exclusive-or (XOR) classification. The solution to this problem was to build an algorithm of multiple Perceptrons, where the outputs of one layer feed into the inputs of another and to

introduce an additive bias term to each neuron (b). These Multilayer Perceptrons are the first example of a neural network, and have been mathematically proven to be able to approximate any continuous function, if the network is sufficiently large [133], [134]. Much as with the regular Perceptron we rely on training the network to set all the internal weights so that we obtain the smallest error between y and $\hat{y}$ for any given input. However, we do not use equation (18) to calculate the change in the weights, instead we take advantage of the chain rule and use two techniques called backpropagation [135] and gradient descent. We first define a cost or objective function such as the mean squared error;

$$E(w,b) = \frac{1}{2}(y(x) - \hat{y}(x,w,b))^2 \tag{19}$$

It is important to note that the input x_i of a neuron is equivalent to the output of the previous neuron $\hat{y}_{i-1}$. This means that to find the error in a neuron, the output of the previous neuron must be known. This is why the technique is called backpropagation, because it begins at the output and work backwards to the input layer, which is more efficient for most ML applications. Calculating the gradient of the error function with respect to each of the input weights tells us how much changing each weight will affect the final error; the gradient will be higher for weights with a greater effect on the overall error.

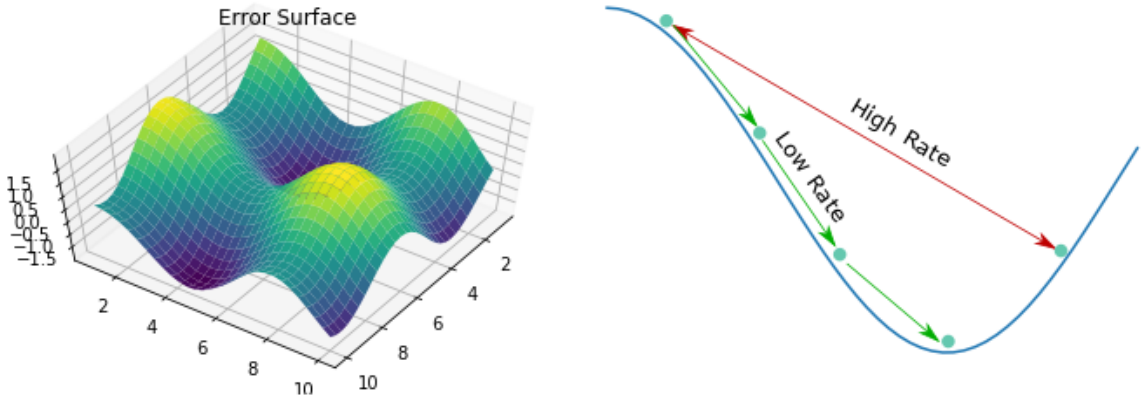

Figure 5.9: Two- and Three-dimensional representations of common issues arising from poor hyperparameter selection during gradient descent.

Hyperparameters are those parts of the algorithm which cannot be learned and must be set by the programmer. These are usually higher-level than the ML parameters, and include things like the number of hidden layers and the learning rate of the model. They are often optimised by the programmer by running various models with different hyperparameters and selecting the best [136]. Two main challenges can arise if the hyperparameters governing the gradient descent algorithm are poorly selected. First, as we can see from the left side of Figure 5.9, there may be many minima within the error surface. We are only interested in the global minimum. Also, as we can see from the image on the right, the rate at which we descend the error surface must be carefully controlled. While large adjustments converge at the minimum faster, there is also a high probability that the algorithm will begin to oscillate about the minimum, never reaching a single value. Of course, a rate which is too low will extend the run time of the program, which may impact its utility. One of the most common methods for avoiding local minima is stochastic gradient descent [137], [138]. Where in regular gradient descent, the gradient is calculated after all the training data has been fed through the algorithm (one epoch), in stochastic gradient descent it is calculated after each training sample. The theory is that the gradient calculated from

each sample will sum to the gradient of the whole dataset, but with small perturbations that will be able to move the algorithm out of local minima. To balance learning speed but also avoid oscillations about the minima, we can use an adaptive learning rate, which reduces how far we move along the gradient after each iteration. Also, to speed gradient descent up, we can use 'momentum', where the previous gradient is factored in to the calculation at each iteration [139], [140]. Error gradients can also be helpful when it comes to understanding the decisions made by an ANN. A common complaint is that the system of weights and biases used to predict an output is incomprehensible to humans. Gradient analysis, a technique available in programs such as KERAS, allows us to add human intuition and guidance back into the process [141]. Gradient analysis transforms the error gradient into information that a researcher can understand and interpret. This way, if attempting to minimize or maximize a certain value, a researcher can guide the algorithm to where they predict this point to be, by analysing which changes bring the algorithm closest to the desired point.

Using these more advanced techniques, a greater variety of problems can be solved. Machine learning tasks are generally broken down into three types, classification, clustering, and regression, summarized visually in Figure 5.10. Classification is the earliest machine learning task, recall the Perceptron, which was a linear classification algorithm. Classification is the process of dividing a dataset into one or more distinct classes. Classification is still one of the more common applications of machine learning, for example identifying traffic signs in autonomous cars [142], or identifying hand written characters, even complex calligraphy [143].

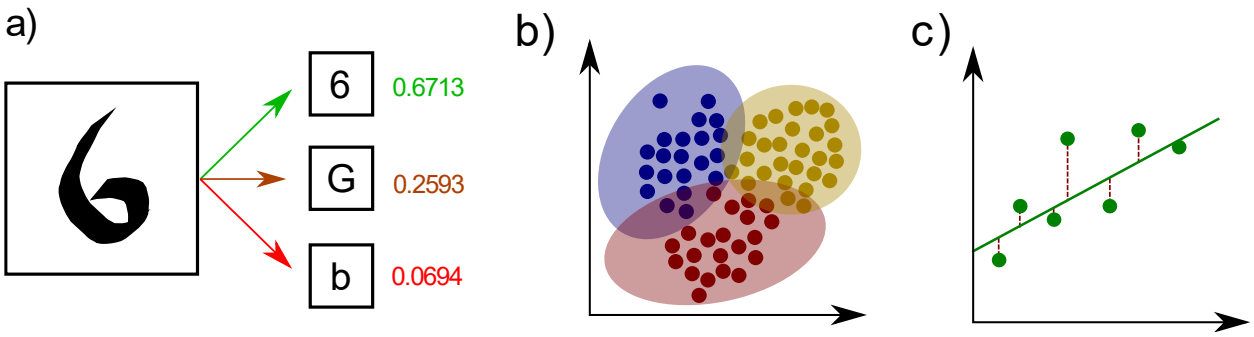

Figure 5.10: Illustrations of the three types of ML problem, a) Classification, b) Clustering, and c) Regression

Clustering is very similar to classification, but instead of dividing the data into predefined classes, the algorithm attempts to find relationships between them independently. For example, separating all the images available online into separate groups relevant to any particular search would be an arduous task, but clustering algorithms are capable of finding similar images and presenting them to the user without assigning specific labels [144]. Regression problems can again be thought of as an evolution of classification algorithms, except in this case instead of discrete groupings, the output of the algorithm are continuous functions. For example, regression models can be used to predict financial performance of stocks and currencies, based on prior observations [145].

One of the most noticeable differences between ML models optimised for these different tasks is in the selection of the activation function in the output layer. All neurons for all sorts of ML models, have an activation function, their purpose is to transform the sum of the neurons' inputs into an output which can be either interpreted by another neuron or as part of the output of the network. Four of the most common activation functions are shown in Figure 5.11. While neurons in the hidden layers of deep neural networks will often use the rectified linear (ReLU) [146] or Leaky ReLU [147] activation functions for efficiency, neurons at the output can use a wide variety of activation functions. Regression problems will most often use a linear activation

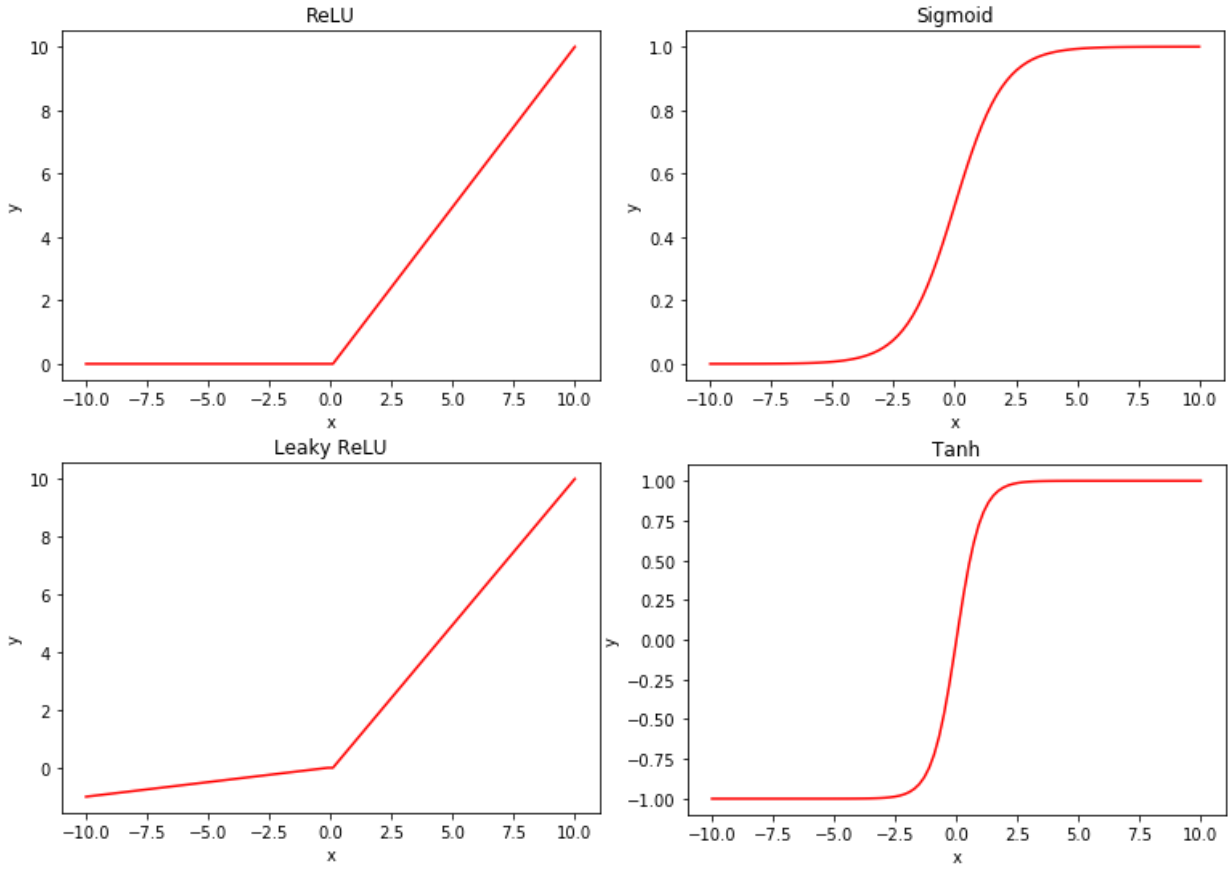

Figure 5.11: Four common nonlinear activation functions

function at the output layer, but must still use nonlinear activation functions in the hidden layers. This is because networks with only linear activations can only ever give linear functions as an output (linear transformations of linear functions will always be themselves linear). Binary classifiers, like the perceptron, will use a simple step function as the activation function. Since the output of a neuron with a step activation will always be either one or zero, it is very good at deciding between two classes (e.g.: 1 = “cat”, 0 = “not cat”). However, when there are more than two classes involved in a problem, more than one neuron may activate at once, and the algorithm is incapable of determining which is the correct output. To correct for these issues, we often use nonlinear activation functions such as the logistic function and its variants, sigmoid, tanh, or softmax when performing more complex classification and clustering tasks. Neurons with

sigmoid and tanh activation functions behave similarly, the difference is that tanh has a steeper gradient than sigmoid [148], but the sigmoid function is bound between $0 \leq y \leq 1$, and so better represents probabilities. For supervised classification problems with a known number of classes, the softmax activation function is also a good choice, as its outputs are also bound between $0 \leq y \leq 1$, but divides each output by the sum of all outputs [149]. Essentially, softmax directly gives us the probability that a given input belongs to all target classes, but where other activation functions give only one output per input softmax gives one output per class for each input. Of course, there are many different activation functions that I have not covered here, it is still an active field of research as almost any differentiable nonlinear function will work, and better alternatives are being found all the time. However, the ones presented here are found as part of most available ML packages, and are therefore still very commonly used.

Figure 5.12 shows a very abstract schematic of an ANN, with the input, output and first/last hidden layers shown. The hidden layers are simply any layer which is not either an input or output; hidden layers are referred to as such because we cannot directly see their inputs and outputs. The number of hidden layers is called the depth of the network, deep neural networks (DNN) generally refer to any network with more than three hidden layers. The number of neurons in each layer is referred to as the width, and increasing either of these values will allow the network to approximate more complex functions. However, since each neuron is connected to all neurons in the layers before and after (referred to as a fully-connected layer) we often prefer to increase the depth of the network and not the width, as wider networks are exponentially more complex.

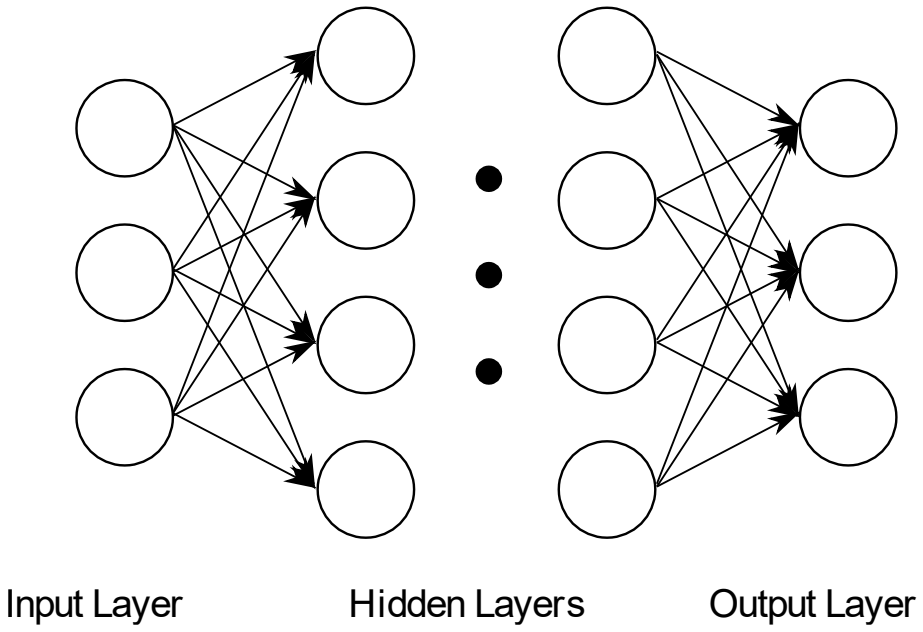

Figure 5.12:Generalized data flow within a neural network

There are many different algorithms in the ML toolbox, many of which are specialised towards solving very specific types of problems. For example, convolutional neural networks which will be discussed in Section 5.2.2 are most often used for image recognition problems, and have specialized features to assist with solving this kind of problem. Most of these techniques fall under the umbrella of deep learning, or DNNs, and training these networks is much more complicated than for a single neuron. Below are some important tricks and definitions which can be used to improve the training of any network, such as bagging, boosting, and transfer learning, and will be referred to throughout the rest of this report.

- **Bagging:** This is short for bootstrap aggregating, and is a method for reducing the variance in the predictions of an ML algorithm. This meta-algorithm allows you to generate additional training data from the original dataset by sampling from it with replacement to create a large number of datasets of the same size and distribution. Training the algorithm on these slightly altered datasets will not make the model better at predicting specific results [150], as the

sampling process is completely random but it all comes from the same initial data. Because the model is trained on the same data through different iterations, but at different stages and with some missing, it gets better at learning the broader trends and not the individual differences between samples.

- **Boosting:** Another meta-algorithm for improving the performance of an ML algorithm, but in this case it is the predictive capability which is improved, and not the variance. Boosting works by training multiple models, each of which may not perform well by itself, but then taking a weighted vote of the outputs of all models. This result is mathematically proven to be more accurate than the individual models used during training, as the error in any one prediction is offset by the error in another [151], however this may not extend to testing error.
- **Convolution:** Is the mathematical operation of multiplying two functions to produce a third which directly expresses how the shape of one is affected by the other. In the case of ANN's, a convolution is usually the multiplication of two matrices, one input and one filter in order to detect specific features within the input. For example, generalizing a square image of 100 pixels/side as a 100x100 matrix, convoluting that image with a filter may allow an algorithm to detect the edges of all objects in that image, or apply a blurring effect [152]. This is the process from which the convolutional neural network derives its name.
- **Cross-Validation:** Is the process of testing the ML algorithm on data that was not used to train it, to ascertain whether it has effectively learned the patterns present in the data, or has just memorized the training input. Commonly the training data will be randomly divided into 'n' sub-sets, and 'n-1' of these are used to train the model, whereas the last is used for

validation. The model is then trained ‘n’ times, so each sub-set gets used for validation one time. Ideally each sub-set will contain a representative distribution of the entire dataset [153].

- **Data Density:** A measure of how well represented an area is within a given dataset. In areas where there is a lot of data, ML algorithms can very accurately interpolate between these data points and their predictions will be stronger. Conversely, where the data is sparser there can be a greater uncertainty in the prediction. Figure 5.13 is a simple 2D representation where the shaded blue areas are the limits of where the ML algorithm predicts the true curve may lie. A low data density is not always negative; when ML is designed with the aim of exploring unknown areas of the data space, areas of high uncertainty are good candidates for further experimentation or simulations [87].

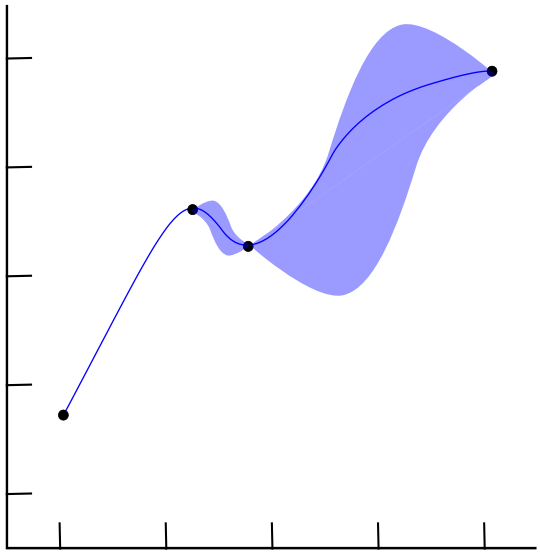

Figure 5.13:Predicted curve with varying data density

- **Kernel:** Commonly referred to as ‘the kernel trick’, this is a way of transforming data using a function (kernel) so that it is linearly separable, but in a higher dimensional space. For example, we want to classify two groups of data, but in two dimensions they are represented by a pair of concentric circles, as illustrated in Figure 5.14. Clearly there is no straight line which can separate these two groups. However, adding in a third dimension can transform the data points so that they can be separated by a linear plane. This method is more efficient than simply trying to find a non-linear separator, and there are many functions which can be used as a kernel, depending on the dataset and problem [154]. Gaussian Process Regression (GPR) is a kind of regression model that uses a kernel to improve its accuracy, as will be discussed in Section 5.2.1.

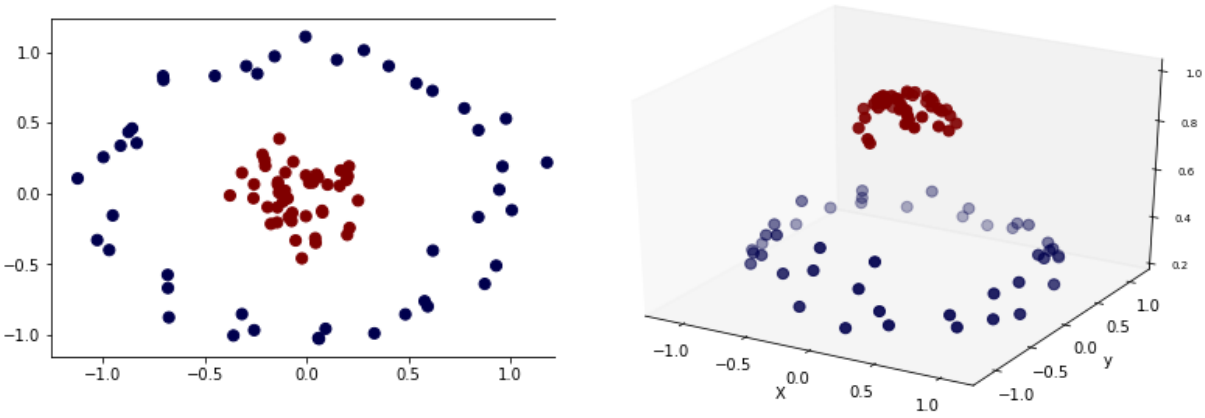

Figure 5.14: The effect of adding extra dimensions on previously non linearly separable data through the kernel trick.:

- **Training:** The overall process of setting the weights and biases of the neurons in a net is referred to as “Training”, but this also refers to a specific phase of that process and the data used to perform it. This occurs in three phases, each using a unique set of data comprised of corelated inputs and outputs. In the first phase a network is given a set of inputs, and its accuracy compared to known outputs is compared using some error function. The weights and biases are adjusted through backpropagation, and it is given another set of inputs. This

phase is referred to as *training*. Next, the model is validated against a portion of that data that was held back and not used for training, this is the *validation* phase. Validation is important, because the model could be memorizing the correlated inputs and outputs during training, so exposing it to new data exposes this fact as the model will provide a very poor result on the new data. At the conclusion of the validation phase, model hyperparameters will be adjusted, and we say that training has completed one *epoch*. Training usually occurs over multiple such epochs, until the validation error is sufficiently low. Once the model hyperparameters, weights, and biases have all been selected a third set of data is introduced. This is the *testing* phase, and the error of the model here is the best indicator of its performance when it will be introduced into a live environment.

- **Transfer Learning:** Is a process for speeding up the training of new ML models by simply re-using an already trained ML model to solve a new problem. For example, in image recognition the algorithm must learn to recognize basic features like edges and shapes before it can begin to classify the images as being of a specific object. Transfer learning is the process of using an algorithm which has already learned these lower-level tasks to extract different but related higher-level data. This saves a significant amount of training time vs creating a new model for each task, as long as the lower-level features are common between the two models [155], [156]. An example of this process is explained in Chapter 7, where I leveraged transfer learning to massively expand the number of material combinations we could successfully measure using our lab's ellipsometer.

5.2.1 Gaussian Process Regression

We have already discussed "classical" regression in Section 5.2, and Gaussian process regression (GPR) is a further advancement of that technique, utilizing a unique probabilistic approach to predict the underlying relation between sets of inputs and outputs. Linear regression, shown on the left of Figure 5.15 assumes that the underlying relation is linear, and finds the best fit by minimizing the error between the line and each of the data points. GPR, shown on the right, instead assumes that all of the data points are distributed with heterogenous noise dependent on *x* via a normal distribution, as opposed to noise with an independent and identically distributed distribution [157].

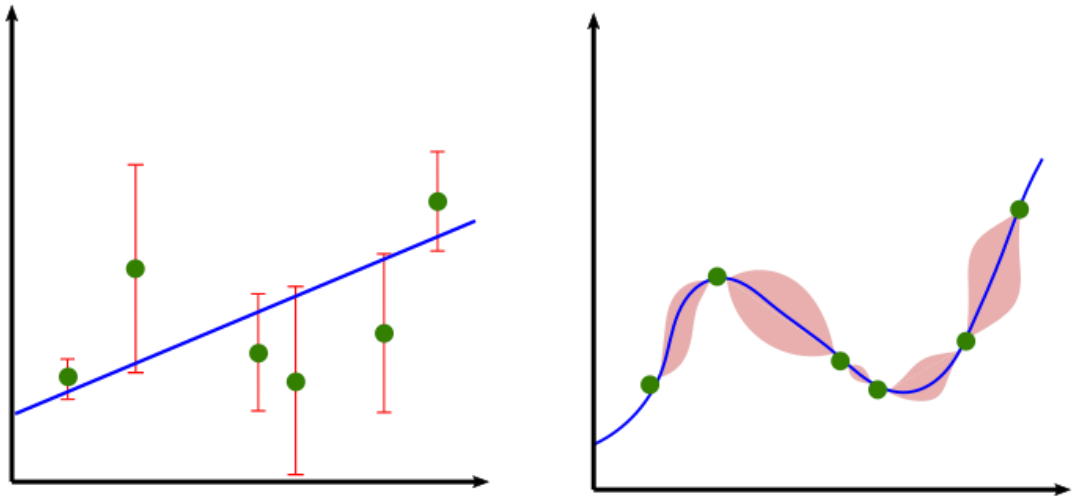

Figure 5.15: Linear regression (left) vs Gaussian process regression (right) over the same arbitrary data points.

This is quite simple to visualize in two dimensions, but real problems usually have an array of dependent and independent variables; the multivariable gaussian function is described by the probability density function:

$$\mathcal{N}(x|\mu,\Sigma) = \frac{1}{(2\pi)^{D/2}|\Sigma|^{1/2}} exp\left[-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right] \quad (20)$$

Where D is the number of dimensions, x is the output variable, μ is a vector representing the mean and Σ is the covariance matrix of dimension D. Covariance is a statistical term which represents the mutual variability of a pair of variables, and the covariance matrix being square means we have the variance of each individual variable down the diagonal, and each variable pair is represented off-diagonal. Simply put, a positive covariance means that the two variables are positively correlated, and a negative covariance means they are inversely correlated, with the magnitude of the covariance being the mean of the variances common between the two. From the covariance matrix, we can start to express some degree of knowledge about the underlying function – the beginning of the probabilistic model underpinning GPR.

The next thing to note about GPR when compared to linear regression, or even nonlinear regression, is that it does not assume that there is *one* function that will accurately describe the relations between the inputs and the outputs. Rather, it assumes that there are *many* such functions, and seeks to find the one that is the most likely to represent the underlying reality. In Figure 5.15 above, we can see shaded areas to the sides of the presumed function line, which represents the range of possible functions. If we were to add a new data point to that line, we would have to update our assumptions about which functions are possible, we have constrained our prior assumptions through the addition of new data. This process holds true even in the case of noisy data, as even random noise is more likely to fall close to the actual function given our initial assumption that the data falls within a normal distribution of said function. We can further tune the probability distribution to better fit any knowledge we have about the form of the function we are trying to find by changing the kernel used to smooth the input data. The radial basis function is commonly used, but packages such as Matlab or KERAS may have additional

kernels available to the operator, and the selection of which kernel to use can be part of the hyperparameter optimization process.

5.2.2 Convolutional Neural Networks

Convolutional neural networks, or CNN's are most commonly used for image recognition tasks, such as handwriting recognition, because they are very good at identifying and extracting important features from large sets of data [152]. The basic function of each layer in a CNN is designed to mimic human sight, in fact CNN's are the artificial system best suited to vision-based tasks which humans perform well such as identifying hazards while driving. Inputs are

$$\underbrace{\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}}_{Input\ Matrix} \quad \underbrace{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}}_{Filter\ Matrix} \quad \underbrace{\begin{bmatrix} 4 & 3 & 2 & 4 \\ 1 & 3 & 2 & 4 \\ 3 & 3 & 3 & 3 \\ 3 & 2 & 2 & 3 \end{bmatrix}}_{Feature\ Map}$$

$$\sum \begin{bmatrix} 1 \times 1 & 1 \times 0 & 0 \times 1 \\ 0 \times 0 & 1 \times 1 & 0 \times 0 \\ 1 \times 1 & 0 \times 0 & 1 \times 1 \end{bmatrix} = 4$$

Figure 5.16: Convoluting an input and filter to find a feature map of a simple binary matrix

given to a CNN in the form of a matrix, often each cell representing the relative strength of red, blue, or green in each pixel of an image.

The first operation performed by a CNN is to convolute the input matrix with another, smaller matrix called a filter. The values within this filter matrix are adjusted during training, and each input is exposed to multiple filters. The purpose of these filters is to extract important data from the input, such as the presence of different shaped curves which may later combine to form a specific letter; this output is called a feature map. A simple example of this is shown in Figure

5.16, here a 3x3 filter is applied to a 6x6 input, moving by either one column or row to fill a 4x4 feature map. Convolution is a linear process, and as discussed in the section on how ANN's

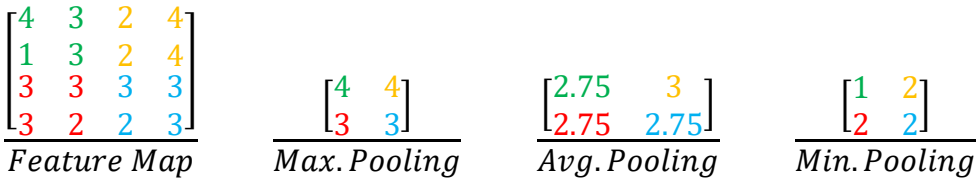

Figure 5.17: Three most common pooling techniques

function, adding nonlinearity vastly increases the number of problems we can solve with an ANN. Accordingly, the ReLU activation function is commonly applied in this convolution layer as it requires relatively little computational power and this layer can often be very large.

The next step in a CNN is to reduce the size of each feature map while retaining as much data as possible. This is performed using a process called pooling, where a selection criterion is applied to subsets of the feature map so that only useful data is passed forward to the next layer. There are three common techniques for pooling data, taking the maximum, average, or minimum of the selected subset. Using the feature map from Figure 5.16, applying a 2x2 pool Figure 5.17 shows how the feature map can be transformed with these three pooling techniques. Not only does this make the input more manageable, it makes the network less likely to overfit the results by trimming unnecessary data, and makes the network more easily able to handle small changes in the input. Older CNN's would misclassify images if the inputs were shifted by as much as a pixel, and this layer helps prevent that. There may be many iterations of convolution and pooling layers, but eventually the outputs of the last pooling layer are fed into a more conventional ANN. This ANN takes the simplified features fed to it by the convolutional and pooling layers and then usually uses softmax to classify the image based on the previously identified features. Real CNN's deal with much larger inputs and have much more complex filters than those shown here,

and the feature maps are correspondingly complex. Observing feature maps, we can often see a hierarchy, where the first convolution may produce uninterpretable filters, but subsequent layers are recognizable as being part of a larger image

While CNN's are most known for commercial image recognition and self-driving cars, the ubiquitous image Captcha™ problems online show what data programmers are looking to work with; it should be no surprise that they have many applications in materials science as well. For example, scientists trained a CNN to predict the toxicity, activity, and solvation free energies of compounds using only 2D images of compounds, outperforming conventional DNN's which had been trained on much more detailed data [158]. One major advantage of CNN's is that they are very good at transfer learning; low-level features such as chemical bond length and orientation can easily be applied to molecules on which the original algorithm was not trained [159]. In both the above papers, the coordinates of molecules were mapped to an 80x80 pixel image and passed into a CNN with three fully-connected layers and 16 convolutional filters. Because the network wasn't given a lot of chemical information, it was possible to observe it learning the features of the molecules which lead to its properties. For example, the CNN learned without prompting that both toxicity and activity are highly dependent on the functional groups present within the molecule. This kind of spontaneous learning is invaluable for potentially complex systems where the interdependencies between parameters may not be known beforehand.

5.2.3 Material Science with ML

Material property prediction is an iterative process, as once the predictions from an ML model have been tested, they can be used to further improve that model until a desired level of

accuracy has been reached. There are four main steps in applying ML techniques to material science;

1. **Feature Engineering:** The process of transforming the desired inputs into a form which can be more easily interpreted by an ML model. For example, representing the bond angles and functional groups of a complex molecule as a series of numbers, or representing the pixels of an image as relative colour ratios. This can also include evaluating the importance of certain features, for example using principal component analysis, and reducing the number of inputs to make the resulting model more efficient. In some cases this can be done manually, or another algorithm can be used.
2. **Model Training:** This is the phase where the dataset created in step one is split into training, test, and validation sets. The training data is fed into one or more ML models with varying hyperparameters until the point where the test dataset can be passed through the algorithm with minimum predictive error. In this phase both the desired input and output are known.
3. **Model Evaluation:** In the model evaluation phase, the validation data set is used to compare the different ML models from step two, and the one with the best predictive properties is selected.
4. **Prediction:** The ML model selected in step three is then given completely new data, consisting of just the input data. If the model has been properly trained, it should be able to predict the desired properties well. These predictions can be verified with further experiments, creating a new, larger dataset which can be used to start the process over again.

Within these four steps there is a lot of variation in the type of features and ML architectures that can be used. TABLE III lists a number of properties found using ML, and the specific types of algorithms used. The goal of this type of algorithm is usually to guide materials scientists, and take some of the guesswork and trial-and-error out of the process of designing specific materials, and careful model design can allow meaningful predictions to be made even if the dataset is incomplete or 'sparse' [160]. MATCALO is an example program, which relates intrinsic material properties, and the changes made by various processing techniques to allow engineers to create materials tailored to specific tasks [161]. This case is quite interesting, because different processing techniques can have either complimentary or negating effects when applied together. Being able to predict the most efficient set of processing steps will save time and money in producing complex alloys like those used in the aerospace industry. Here regression decision trees were used as the ML architecture, where each leaf in the tree represents the value of the predicted property if the processing steps follow from the root of the tree to that leaf. The algorithm used then works in reverse to a normal decision tree, starting at the leaves with desired properties, before selecting the best roots. Applying restrictions and desired property values to the system forces the algorithm to select only the best candidates, for example it may be desirable to minimize the processing temperature due to some limitation of the intended use of the material. This approach may be useful when trying to use an ANN to optimize the growth conditions of our thin films.

TABLE III: MATERIAL PROPERTY PREDICTION AND CHARACTERIZATION USING MACHINE LEARNING

Property of Interest	Algorithm	References
Nanomaterial Solubility	Regression	[162]
CNT/Nanofibre Morphology	ANN	[163], [164]
Magnetic Property	ANN	[165]
Toxicity	Decision Tree/Random Forest/Support Vector Machine (SVM)	[166]–[168]
Biological Activity	ANN/Linear Regression	[169]
Catalytic Activity	ANN/Bayesian Linear Regression	[170]–[172]
Nanorod Dye Adsorption	ANN	[173]
Electron Transfer Properties	ANN/K-Means /Random Forest/Regression	[174]
Reactivity	ANN/Recurrent Neural Network (RNN)	[175]–[177]
3D/Crystal Structure	ANN/Deep Neural Network (DNN)/Regression	[178], [179]
Grain Boundaries/Defect & Phase Evolution	Convolutional Neural network (CNN)/Clustering	[97], [180]
Raman Spectra	Clustering/Principal Component Analysis (PCA)/Regression	[181]–[183]
Quantum Dot Charge/State	CNN/DNN	[184]
Band Gap & Alignment	ANN/DNN/SVM	[185]
Perovskite Lattice Constants	ANN/Random Forest/Regression	[186]–[188]
High Entropy Alloy Phase	ANN/K Nearest-Neighbours (KNN)/SVM	[189]
Nanoparticle Lattice Orientation	CNN	[190]
Graphene Mechanical Properties	ANN/Decision Tree/KNN/SVM	[191]
Structure-Property Relationships for Hexagonal Boron Nitride and Graphene	ANN/CNN	[192]
Nanoparticle Formation	Regression/SVM	[193]
Dielectric Breakdown	Regression	[194]

As mentioned previously, discovering new materials is most often a classification task, but that is not the entire picture. These algorithms often consist of more than one ML algorithm working in tandem, one acting as a predictor to find what the structure or property of a material will be given a set of inputs, and another to classify that material as part of the desired group or not. The work done by Wen et. al. in [87] is a good example of exploring a wide range of data to find new materials. In this paper the authors were searching for new high-entropy Al-Co-Cr-Cu-Fe-Ni alloys, eventually finding a new alloy with 10% greater hardness than any used in the training data. One advantage they used in their model was online learning, where the hypothesized alloy from each iteration of the ML model was synthesized and its measured hardness was added to the training data for the next iteration. This is a good way to fill areas of the dataset which are sparse. The ML algorithm then predicted both the hardness of new alloys and a utility function which selected the best candidate for further experimentation. This decision was based on the predicted hardness, and would prefer experiments in less-explored regions of the data. When using more property data, they found which properties are highly correlated, and removed the ones with higher error, streamlining the model and increasing accuracy while decreasing the time to get a result. To streamline the model further, they compared models which were trained not to use all available features, and found the three most important features. This paper is particularly relevant to my research because it highlights some common pitfalls in materials discovery, and ways to avoid them. While I will likely be using fewer materials to compose the new thin film materials I aim to discover, the greater number of processing variables means I must take care when selecting features to feed into the model. Also, the idea of combining multiple models together to find the materials most statistically likely to have desired properties will be very helpful to avoid the ‘black box’ problem often associated with ML

models, where the model gives accurate results but the process of obtaining those results is unintelligible to an observer.

ML techniques can also be used to assist materials researchers in other ways, such as de-noising measurements, making them less sensitive to initial guesses [195] and extracting extra information vs traditional techniques for the same data [196]. In these two sources they used ML to improve signal-to-noise ratio of AFM, but have done the same with SPM [197] and STM data [196]. Using ML in tandem with conventional data-collection techniques also captures fast/dynamic information, which is hard to do with conventional filtering. ML can also be used to automate tasks which are trivial but time consuming for humans, such as grain edge detection in perovskite materials [198], where the researchers used classifiers to detect boundaries in the same manner as the human eye. Finally, ML can be used to speed up ab-initio simulations [199] while retaining accuracy by identifying data the simulation produces but is not needed, and extracting only relevant data. Alternatively, instead of speeding up a single simulation, ML can be used to cut down the number of simulations needed [200]. The authors trained an ML algorithm on simulation data, and used Bayesian inference to determine the accuracy of ML predictions. If the predictions were accurate, the results of the ML model were used, saving time over simulating all necessary datasets.

6 Neural Network Assisted In-Situ Optical Characterization and Feedback Control

6.1 Introduction

AP-SALD is in the early stages of commercialization, but critically it still lacks process control capable of leveraging its unique ability to form oxides of such varied compositions. Optical measurement techniques, such as reflectance spectrometry, are ideal for incorporating into commercial processes because they are capable of making rapid measurements without damaging the sample being measured. As discussed in Chapter 2, in-situ optical characterization was already implemented into the AP-SALD system by another PhD student, Kissan Mistry. Using a reflectance spectrometer, the thickness, optical band gap, transparency, and changes in surface roughness of the film can be found by fitting a model to the measured spectrum. In Section 5.1.1 we briefly discussed other researchers work on using ML to fit reflectance, focused on industrial applications, but these are more recent advancements. Conventionally a naive optimization approach like a grid search might be used as a means to narrow in on the correct fitting parameters, rather than exhaustively searching through all possible combinations. A more robust means to narrow in on the optimal optical model parameters would be to use a curve-fitting model like the Levenberg-Marquardt algorithm. Unlike the grid search method, this is an iterative algorithm, meaning it is run repeatedly in an effort to narrow in on the most accurate values, but this does mean that it can get stuck in local error minima. The first step with this algorithm is to provide an initial estimate for each parameter of interest, thickness and refractive index for example, as well as pairs of independent and dependent variables. This might be a measured spectrum (independent), and what the spectrum would look like given the specific

parameters fed into the model. Over each iteration, the algorithm adjusts the values of the parameters of interest in order to lower the sum of the square of the deviance between the dependent and independent variables. The Levenberg-Marquardt algorithm is the method most commonly used in commercial optical model fitting software, as it is generally very good at finding the parameters with the least resulting error as long as the initial parameter estimates are close to the actual values. Despite its iterative nature, it is much more efficient than a grid search algorithm for problems with many different parameters.

It has been shown that the properties of thin oxide films grown by AP-SALD are different to those produced using more conventional ALD and CVD techniques. For instance, it has been shown that the presence of atmospheric oxygen during depositions caused increased grain-boundary scattering [201], [202], consistent with our results from Chapter 4 [31]. The ways in which AP-SALD films differ from those produced with conventional ALD is a unique opportunity to create films tailored for specific applications, but complicates the process of creating and fitting an optical model to the reflectance data, requiring a more flexible fitting method with fewer limiting assumptions. As discussed in section 5.1, most non-destructive measurement techniques rely on mathematical models where various film properties are already known, and it is our objective to automate this process. While it would be tempting to use a Levenberg-Marquardt algorithm to fit each and every parameter of an optical model in turn, this is an incredibly computationally expensive process. Since this program needed to be built in Matlab in order to integrate with our control system, any such design would be complex to build, and still run too slowly to produce accurate in-situ results given that the AP-SALD must pause for the length of any measurement, diverging from ideal deposition conditions.

Rather than using previously-reported material properties, and finding the thickness through an optical model, we chose to invert this structure with the help of a database of known films. This database contains the complex refractive index, thickness, and full reflectance spectrum of thousands of films calculated using the Tauc-Lorentz optical model. The Tauc-Lorentz model was selected, as it is a common model for describing the response of semiconductor films to light, as described in Section 5.1.3. By using a combination of grid search and ML methods to estimate the thickness of the film based on the raw reflectance data and the number of AP-SALD deposition cycles, we were able to use a Levenberg-Marquardt fitting algorithm to match our measured spectra with those in the database, as illustrated in Figure 6.1. Having matched our measured spectrum to one with a known refractive index and thickness, we could then use a Tauc plot to track the evolution of a films optical band gap in real-time. By leveraging the power of ML to provide accurate thickness estimates, we were able to produce real-time measurements of both the films thickness, complex refractive index, and band gap, faster and more accurately than with our previously-implemented conventional grid-search algorithms. Being able to monitor the evolution of a film's properties during deposition enabled greater process control, including observation of the nucleation period, nanoscale thickness control, and growth rate control. These facilitate tailoring films towards specific applications. While this initial study was limited to ZnO films, this material has a broad range of applications as a semiconductor material. Being able to match its bandgap to adjacent films in a stack and precisely control its thickness will result in more efficient devices, critical in the age of increased anxiety surrounding climate change and the need for greener technology.

6.2 Experimental Methods and Materials

6.2.1 Film Deposition and Characterization Techniques

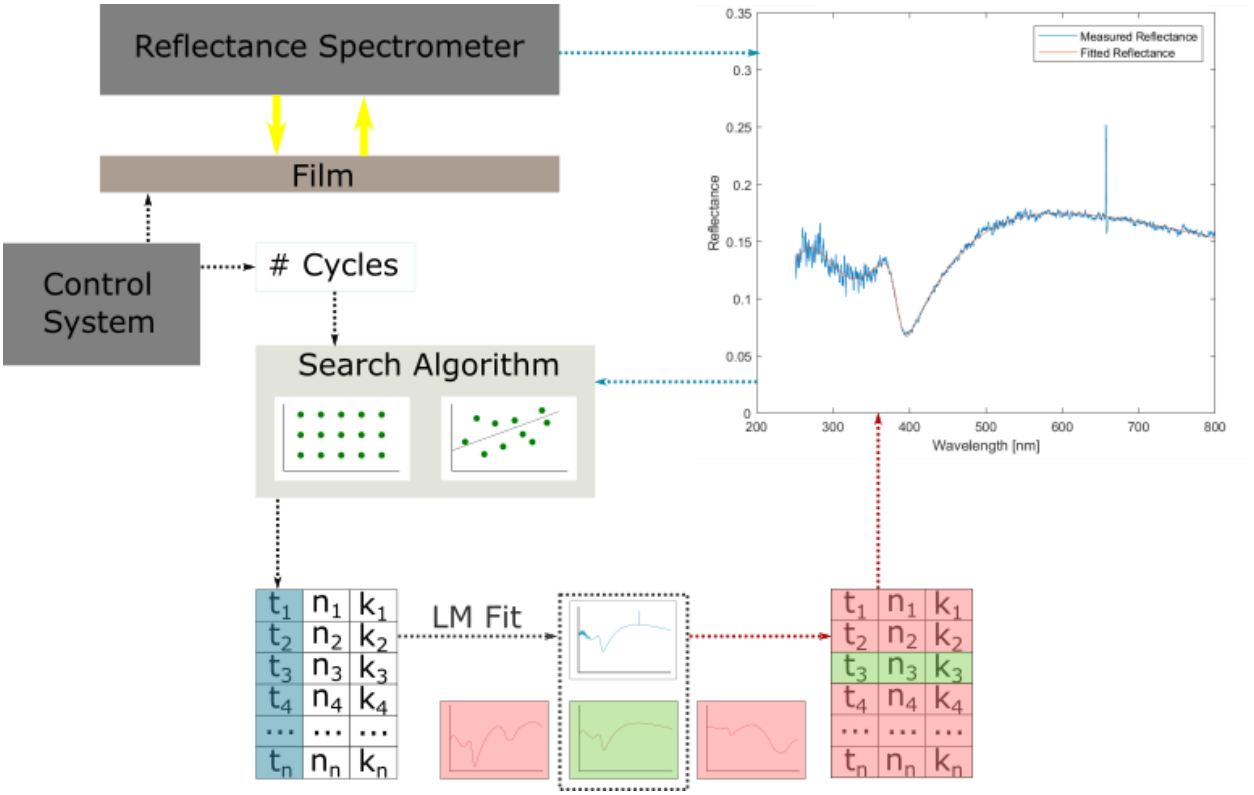

Figure 6.1: Workflow schematic between the reflectometer, search algorithm, and film database.

The AP-SALD process is detailed in Chapter 2, but for the films here we limited our analysis to ZnO films produced between 100ºC and 200ºC, controlling the flow rate by adjusting the relative proportion of gas flowing through the diethylzinc bubbler before reaching the reactor. These conditions are representative of the most common depositions we perform on our system, and as discussed in Chapter 2, still represent a diversity of film properties. We used an Ocean Optics HDX UV-Visible spectrometer, coupled with a DH200 light source and QR600 bifurcated probe to acquire the spectra. The probe was mounted perpendicular to the film surface, and the deposition was made to pause for the duration of the measurement time, a fraction of a second, to ensure that the same spot was measured consistently. The spectrometer

was calibrated using a bare silicon wafer, and the borosilicate glass substrate was also measured to ensure proper differentiation between it and the film during measurement. While allowing the precursor flow to continue has an effect on film uniformity in the direction of oscillation, as the precursors continue to flow during this time, the response time on the valves and MFC's is too slow to ensure that the deposition is not interrupted. This does not affect the area being measured, but would limit the utility of our approach when depositing films for device stacks, in addition it would result in an increased annealing time for the film, altering its properties. This fact further reinforces the importance of measurement speed, one of the major focuses of this work. The measured spectrum has a lot of noise, as in order to make the measurements fast enough to be used in-situ, we have to compromise on the length of time for which each measurement is taken.

6.2.2 Reflectance Measurement Methods

Figure 6.1 is a process flow diagram, illustrating how two spectra are compared, one (blue) is the measured spectrum, and the other (red) comes from a database of films with known thicknesses and optical constants, computed using the Tauc-Lorentz optical model. There were three iterations of the search algorithm component, each of which was designed to provide an initial thickness estimate based on the measured reflectance spectrum and the number of AP-SALD cycles. The results of each search algorithm were then passed to a Levenberg-Marquardt (LM) fitting algorithm to refine the thickness estimate and improve the goodness of fit between the measured and modelled spectra. The post-LM fit thicknesses from each iteration were used to train the next.

The first version of our search algorithm, which was retained in some form throughout the development of the project was a simple grid-search algorithm. This algorithm functions by

choosing evenly spaced values along each axis of the data, and evaluating the outputs at that point, represented in Figure 6.1 in a 2-D plane. The combination of inputs with the output closest to the desired or real value is then selected. This algorithm is useful because it searches across all possible combinations without potential bias from the end user, but with more input axes the time to compute the best combination gets longer and longer, limiting its utility. The grid search algorithm we used outputs an estimate of the thickness of the film in nanometres, using the number of cycles as the initial input, searching a domain that includes thicker films with each measurement. The second iteration was a simple feed-forward neural network with a single hidden layer. This algorithm used both the number of cycles and the raw reflectance spectra as inputs, passing through the hidden layer and providing a thickness estimate. Because this is a very shallow network, three different training methods were tested to ensure the lowest MSE value, the Levenberg-Marquardt, Bayesian Regularization, and Scaled Conjugate Gradient methods, which were the three default options in Matlab. Finally, due to the high degree of noise in our measurements, we used a Gaussian Process Regression model to predict the thickness of the film. Because we know that thickness is a smooth function with respect to cycle number, the underlying assumptions of the GPR model made it an attractive solution to the problem of noisy training data, given that the outliers will be used to train the model without assuming that they are an accurate representation of the films thickness – it allows for varying levels of noise at different values of x unlike most other regression methods.

6.2.3 Machine Learning Methods

Proper treatment of the data is integral to getting meaningful results from an ANN of any type. In this work, the input data was an array of raw reflectance data, the number of cycles, and the desired output is a guess for the film thickness, d. $R(\lambda)$ is the full reflected spectrum with

1,878 individual data points for every measurement, corresponding to a wavelength range of 250 to 800 nm, and can include a lot of meaningless noise. Initially I used all 1,878 inputs, training ANN's of a specific width three times, once with each training algorithm. In most cases this led to a model which stopped improving over subsequent training iterations at poor training MSE values, and in the cases where a model was successfully fit, the training R^2 and MSE were both high. This means that the general shape of the predicted function may be accurate, but there was a large variation in the individual predictions. Because the large number of inputs led to long training times and poor ANN performance, it was decided to pool the input data, averaging each block of six values, reducing the noise. Unfortunately, the amount of data we had available was limited by the initial challenges of the grid search model. If we attempted to train any kind of neural network on data that is mostly noise, we could have no confidence in its accuracy. Therefore, we were limited to using only those films for which we had reasonable confidence in at least some of the thickness data. For the first-generation ANN, this dataset was comprised of the results of 20 films deposited under different conditions, measured every five cycles. This was supplanted with another 30 films for training the GPR model, as that first-generation ANN provided a significant boost in our program's accuracy. As previously mentioned, the initial thickness estimate provided by the search algorithm is fed into a Levenberg-Marquardt (LM) fitting algorithm, produced using the *lsqcurvefit* function within Matlab. This algorithm iteratively adjusts the thickness, searching through a database comprised of spectra produced using the Tauc-Lorentz optical model until one is found that matches the measured spectrum. The LM algorithm is very sensitive to the initial estimate, or else it may easily fall into a local error minimum and give a poor fit. Iterating over only the thickness is far faster than attempting to fit all of the Tauc-Lorentz parameters individually, but further reinforces the necessity of an

accurate initial thickness estimate. This database also includes the index of refraction and absorption coefficient for the films, and from those values and the thickness the band gap can be calculated using a Tauc plot – given that we know ZnO has a direct band gap. The search algorithm only finds the initial thickness estimate for the LM-fitting algorithm to refine, as this algorithm is very sensitive to initial conditions. Thickness is the ideal sorting parameter because n and k vary by wavelength, but thickness is unique per film. It also has a comparatively greater effect on the shape of the spectrum, as we are only looking at ZnO here. The range of thicknesses is greater than the range of the optical constants, as it is not bound to any range save that of what can be practically grown in a single deposition, whereas optical parameters are constrained by the properties of the material itself. As shown in Figure 6.1 we select a film from the database (red) based on how well it matches the measured spectrum (blue) based on the LM algorithm, and report the goodness of fit (GOF) between the two curves. A basic measure of the total error at each point, a GOF of 1 would indicate a perfect match between the two spectra and allows us to monitor how well our program is predicting the thickness of that film.

6.3 Results and Discussion

6.3.1 Grid Search Algorithm

All versions of the reflectance fitting program use the number of AP-SALD oscillations passed as an initial estimated thickness, serving as a baseline for the search algorithm on the films searched within the database of ZnO reflectance spectra. This was particularly important for the grid search algorithm. Having a baseline means we can move the grid to search in areas of the database more likely to contain the actual thickness. For example, it is not worth searching for spectra corresponding to films 100's of nanometres thick if only 5 AP-SALD oscillations

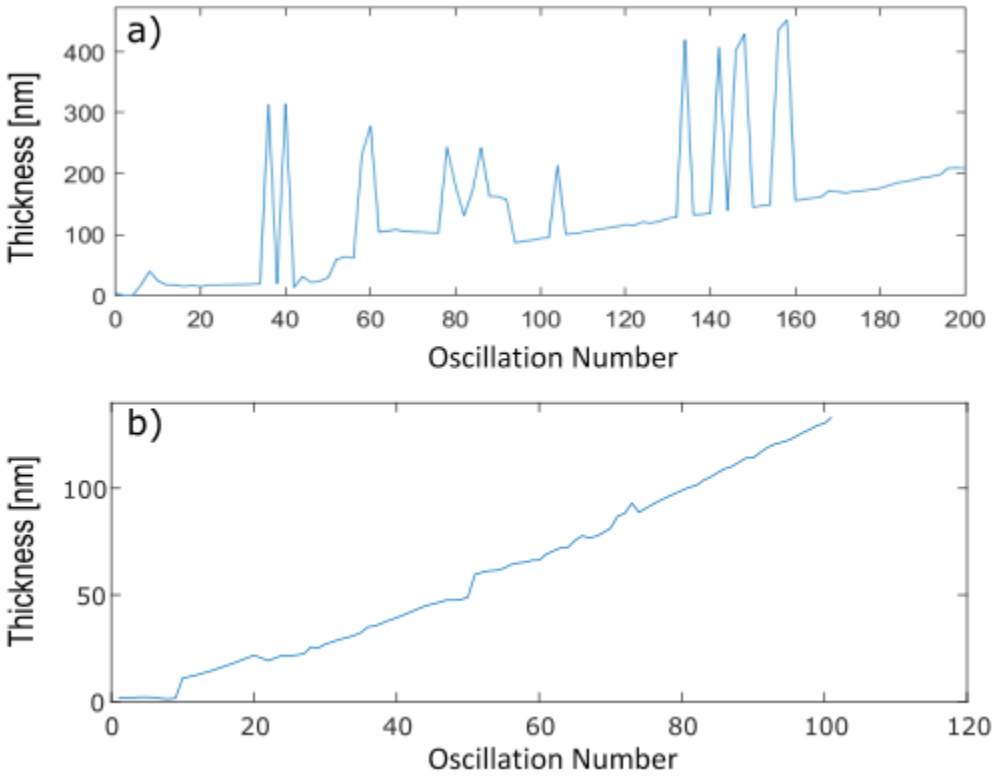

Figure 6.2: Estimated film thicknesses based on a grid search around an assumed growth rate of 1 nm/oscillation, either a) unconstrained, or b) constrained to within ±5nm for the first 30nm, and then to within ±50nm.

have occurred. Figure 6.2 shows the estimated thickness throughout the deposition of a ZnO film at a temperature of 100°C. It demonstrates the difference between a naïve grid search (Figure 6.2a), and one where the range of thicknesses searched was limited (Figure 6.2b), first to within

±5nm of the presumed 1nm/oscillation growth rate until the film was 30nm thick, and then to within ±50nm. Since the naïve version searched the database of ZnO reflectance spectra unconstrained, simply finding the fitted spectrum with the best goodness of fit with no regard for realistic growth dynamics, there were large spikes in thickness where the LM fitting found a higher goodness of fit at an unrealistic thickness, as shown in Figure 6.2a. This lead to the implementation of a grid-search algorithm as in Figure 6.2b where the thickness is searched to only within a specific range to eliminate large thickness spikes of several hundred nm. The growth predicted in Figure 6.2b is much better, but still shows several unrealistic thickness measurements:

- The nucleation period is followed by an abrupt increase in thickness at oscillation 10.
- Several instances of a reduction in thickness.
- Multiple other instances of small spikes in thickness, often followed by a plateau or drop.

Further refinements of the grid search algorithm allowed us to begin measuring film properties through the deposition process with reasonable accuracy. This version of the grid search algorithm used 0.2*oscillation number as the lower bound, and 5*oscillation number as the upper. In addition, a “boxcar” technique was applied to the measured reflectance data, where instead of reporting a single reflectance value at each wavelength, a moving average is taken. Here the width of the boxcar was five measurements, meaning that at each wavelength the average of the reflectance at that wavelength and the reflectance at two nm above and below was reported. This has the effect of smoothing the reflectance data, but if the boxcar is too wide the details of the plot can be erased. Figure 6.3 shows how much better this version of the grid search algorithm performed compared to earlier attempts. This algorithm allowed for measurements of the films thickness and band gap that were very accurate under linear growth

conditions, but still did not handle nonlinearities well.

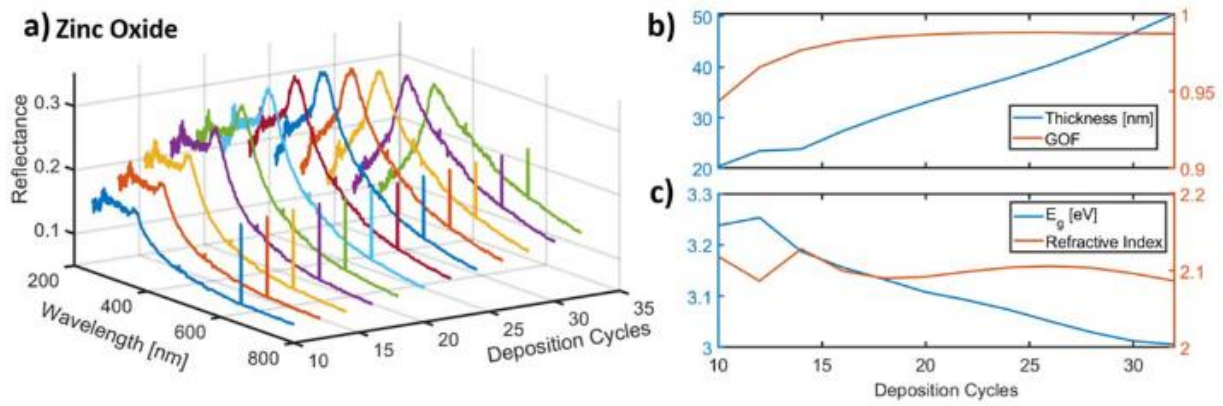

Figure 6.3: a) Reflectance spectrum changes with film growth up to 30 oscillations, b) changes in measured thickness and goodness of fit, and c) evolution of band gap and real refractive index over the same interval – Source: Kissan Mistry [https://iopscience.iop.org/article/10.1088/2632-959X/ab976c]

Figure 6.3b and c show how this real-time data monitoring allowed us to accurately track film properties beyond the thickness under those linear growth conditions. However, the use of thickness bounds restricts the range of deposition conditions under which the algorithm can function (i.e., changes in the flow rate, temperature, or deposition errors which change the deposition rate may not be accounted for). The comparatively low GOF values at low thicknesses show the models struggle to represent the nucleation stage accurately, when growth has not yet reached the point where a monolayer is formed each oscillation. All our attempts still showed that initial "jump" once the film was thick enough, as shown in Figure 6.4, which results in a poor GOF value. Accordingly, these measurements were omitted from the data reported in our paper and Figure 6.3. This is a non-physical nonlinearity, but the measurement technique we described does hold true for films of sufficient thickness.

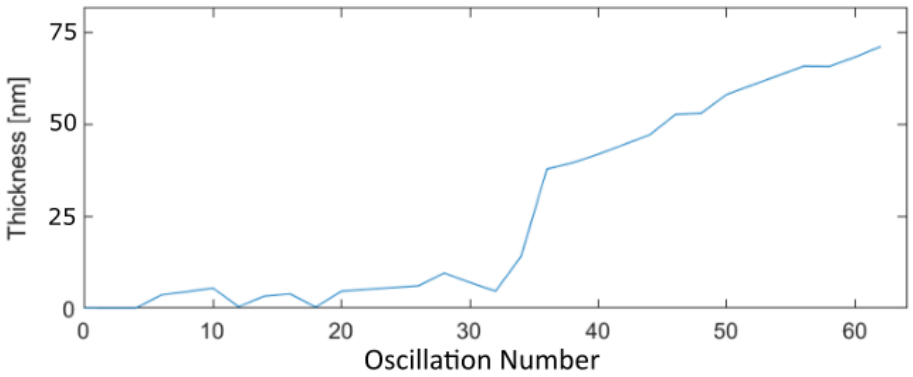

Figure 6.4: Nonlinearity after the nucleation phase still visible, even with the most advanced grid-search algorithm we developed.

6.3.2 Feed-Forward Artificial Neural Network

We noticed a consistent relationship between the shape of the reflectance spectrum and the thickness, measured post-deposition using AFM, as shown in Figure 6.3a. It is this relationship that inspired us to incorporate ML methods. Visible patterns in data are a good sign that there is a mathematical relation underlying it – the relation that the optical model seeks to replicate. By allowing ML methods to find the initial thickness estimate, and combining it with the LM fitting algorithm, it was hypothesized that we could retain the accuracy of the improved grid search method but under nonlinear growth conditions. The first version of the ML model used a shallow neural network, with a single hidden layer between the inputs and outputs. This hidden layer can have any specified number of neurons inside it, its "width", created using the Matlab tool *nftool*. Theoretically, an infinitely wide network can model any regression problem, but the computational cost increases massively the wider a network gets, which is why most research is done using "deeper" networks, with multiple hidden layers in series. The strength of the prediction provided by *nftool* is indicated by two statistical measures, the coefficient of determination 'R^2', and the mean squared error (MSE). R^2 measures how well explained the variance of the output is by the variance of the input in a linear regression model such as the one

used here, and the best possible value of R^2 is one. R^2 is a good statistical measure because it is easily interpretable and can be compared easily between models. An R^2 value of 0.75 means that 75% of the variance of the output is explained by the model, and an R^2 of 0.99 corresponds to 99% explainable variance. MSE, which has an optimal value of zero, is found by calculating the error between the prediction provided by the neural network being trained, and its actual value from the training dataset, squaring it, and averaging this value for all outputs. It is not a measure of the variance of the input and output parameters, but instead the accuracy of the results as produced by a given model. The scale of the MSE will depend on the scale of the variable being measured, as it is based on the absolute difference between measured and predicted values of that variable. We trained the feed-forward ANN using three training algorithms, Levenberg-Marquardt, Bayesian Regularization (BR), and Scaled Conjugate Gradient (SCG) backpropagation using networks of varying width – the results of which are summarized in TABLE IV. In this table, based on the high R^2 values and testing MSE, we can see that many of the models overfit the data, memorizing the output rather than learning the relationship.

TABLE IV: COMPARISON OF THE PERFORMANCE OF FEED-FORWARD ANN'S OF VARYING WIDTH, TRAINED USING LEVENBERG-MARQUARDT, BAYESIAN REGULARIZATION, AND SCALED CONJUGATE GRADIENT ALGORITHMS.

Width	Training Algorithm	Training MSE (nm^2)	Training R^2	Validation MSE (nm^2)	Validation R^2	Testing MSE (nm^2)	Testing R^2
5	LM	0.034	0.999	43.29	0.9951	32.15	0.9976
	BR	0.000001	0.999			27.13	0.994
	SCG	35.77	0.9969	48.44	0.9953	64.60	0.9921
10	LM	0.397	0.99	35	0.99	31	0.99
	BR	2.649	0.9998			11.57	0.9989
	SCG	4.428	0.9996	12.97	0.9987	24.69	0.9977
15	LM	0.059	0.999	32.8	0.997	46.14	0.996
	BR	0.0244	0.999			0.9479	0.999
	SCG	23.25	0.998	48.52	0.995	37.26	0.997
20	LM	$6x10^{-27}$	1	37	0.9957	23	0.9984
	BR	1	0.999			12	0.998
	SCG	18	0.999	95	0.983	156	0.985
25	LM	$9x10^{-27}$	0.999	65	0.995	41	0.995
	BR	1.3	0.999			6.9	0.998
	SCG	160	0.985	247	0.981	279	0.976
50	LM	32	0.997	182	0.992	267	0.9708
	BR	0.00002	0.999			6	0.999
	SCG	294	0.974	498	0.956	367	0.976

Levenberg-Marquardt is a common optimization technique, and is faster than the other two, however, it is comparatively inefficient and takes a lot of processing power. Bayesian Regularization starts with an LM process, but then also seeks to minimize the errors and weights within the system. This means that models trained using this algorithm tend to generalize well,

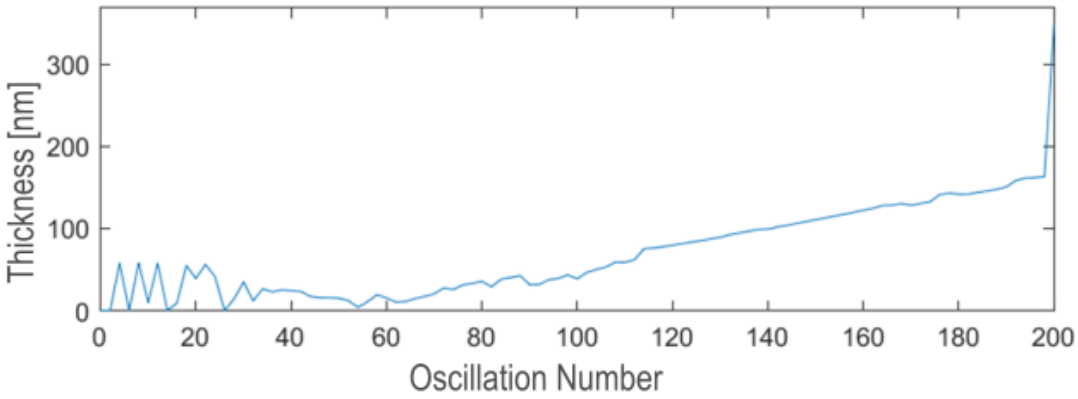

Figure 6.5:Example of fitted film thickness deposited at 150ºC using first generation ANN, with significantly smoother fitted thickness values above 60 nm compared to the model with no ML-assisted fitting

and are less prone to overfitting on small or noisy datasets. Scaled Conjugate Gradient is good for larger problems, as gradients are a more memory-efficient calculation than solving the Jacobian matrices required for the prior two methods. From all this, we expected BR to perform the best, as our dataset was both small and noisy. Comparing the MSE and R^2 values for all three algorithms trained across varied widths, we saw that the 15-neuron-wide network outperformed all the others when trained with BR. The algorithms that overfit on the training data are obvious from this as well; this can be seen wherever the validation and testing MSE are significantly higher than the training value, although R^2 remains fairly close to one for all trained models. For instance, if we look at the 20-neuron-wide network trained using the Levenberg-Marquardt algorithm, we can see that the training MSE is an incredibly low $6x10^{-27}$ nm^2, nearly thirty orders of magnitude lower than the testing MSE of 23 nm^2. We also see how the width of the network can have an adverse effect on the accuracy, as the MSE for both LM and SCG-trained models balloons when the network is more than 20 neurons wide. This is partially due to the nature of

our dataset, as these algorithms are suboptimal for small and noisy sets of data. As can be seen in Figure 6.5, the first-generation ANN was not significantly better than the grid search algorithm, in large part because it relied on thickness data supplied by that algorithm – garbage in, garbage out. Notably this iteration of the search algorithm did not show zero film growth in the nucleation phase, instead showing wildly oscillating thickness data. This shows that it "knows" that the thickness is nonzero, but because it was not supplied with good training data in this region it was unable to provide a reasonable estimate either. There is promise in this method, but it is currently unconstrained, similar to the first-generation grid search.

As shown in Figure 6.6 the first-generation ANN performed worse than the grid-search

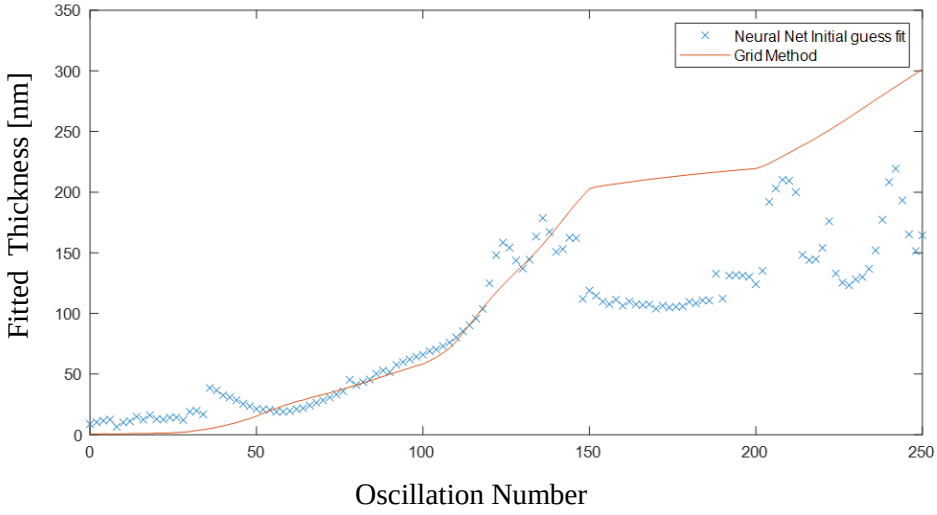

Figure 6.6: Plot of film thickness deposited at 150ºC vs oscillation number, comparing ML-assisted fitting and basic grid-search methods at high oscillations – image courtesy of Kissan Mistry

algorithm at high thicknesses too. The initial training dataset did not include films of greater than 150nm thickness, and so the ANN was incapable of extrapolating beyond this value. NN models aren't good at extrapolation in general, it is an inherent part of how they operate – data must be selected to encompass the full range of values for the target variable. Since we already have good results from the grid-search algorithm outside the nucleation phase, we decided to limit the next version of the ANN to measuring films from 0-50nm before switching to the grid-search method,

eliminating the future problem of needing larger and larger datasets to encompass even thicker films. Grid-search as an algorithm is not bound by thickness, so its upper limit is bound only by our database of ZnO films, not by its own method of operation. We still bind the domain in which the grid-search operates in order to speed up the calculation, as per Section 6.3.1, but there are further improvements we can make to the structure of the ML model itself.

6.3.3 Gaussian Process Regression

As discussed in Section 5.2.1, Gaussian Process Regression (GPR) is designed to work with noise that varies with respect to the independent variable, predicting a centreline based on the training data and assuming that any predictions which result in an outlier are less likely than those falling close to said line . This is a *probabilistic* binding of the predictions, but is not a hard limit, so it can find a trend that isn't visible from the raw data, and makes this kind of algorithm better suited for working within the limits of our dataset. Figure 6.7a is an example of how a GPR algorithm is capable of finding a relationship underlying noisy data, and Figure 6.7b shows the training histogram for the algorithm presented herein.

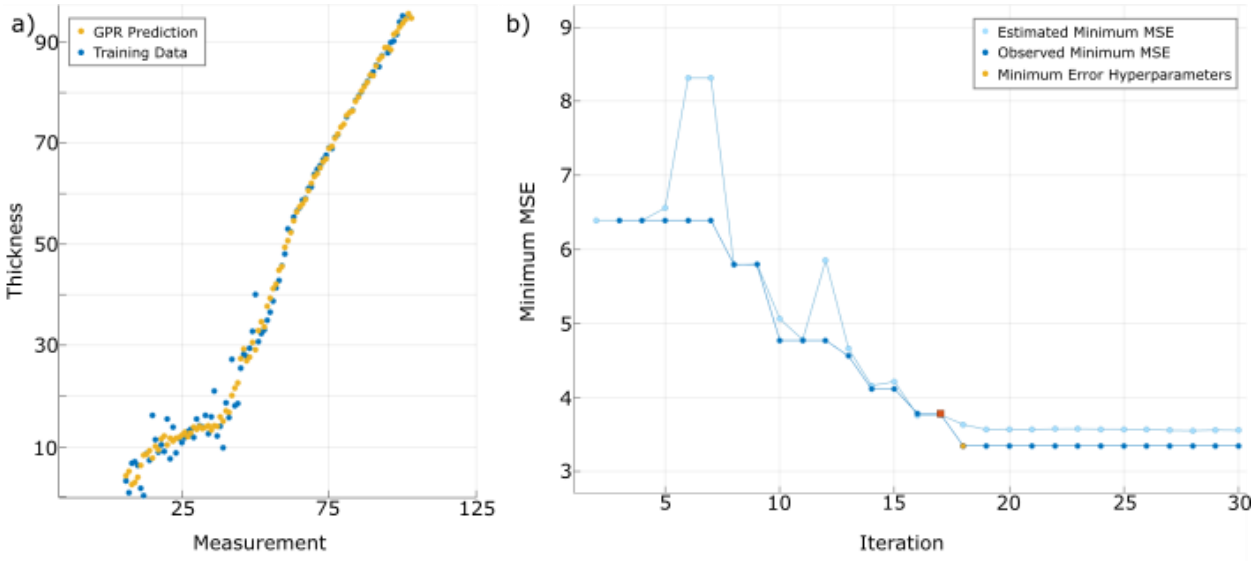

Figure 6.7: a) Actual (blue) vs predicted (yellow) values for reflectance data during training of the v2 GPR model, noise reduction at low thicknesses is visible. b) Training histogram for the v2 GPR model, iteration with lowest observed MSE indicated in yellow

The testing MSE for the GPR model was 3.397 nm^2, as can be seen on the training histogram in Figure 6.7b. This is the iteration of the algorithm that had the best performance measured by MSE, although subsequent iterations had similar values. An MSE of 3.97 nm^2 is higher than the MSE of the v1 model, which seems worse. However, Figure 6.7a shows us what is really happening; because the GPR model explicitly searches for a smooth line that explains any noise, it was able to better approximate the real thickness values at the expense of having higher error between the measured and predicted variable. The reason the MSE was so low for the v1 ANN was that it learned to replicate the noise present in the grid-search-provided data, and therefore it also provided unreliable thickness measurements. This is a disadvantage of algorithms that are designed to approach data with no presuppositions about it – those suppositions may be an integral part of understanding the reality behind the data. Once the thickness reached a certain value, because the training data was more consistently accurate, the GPR model successfully converged to the results of the grid search algorithm. Since the data at this point was not noisy, there was no advantage to using the GPR algorithm, and as discussed earlier, the grid search

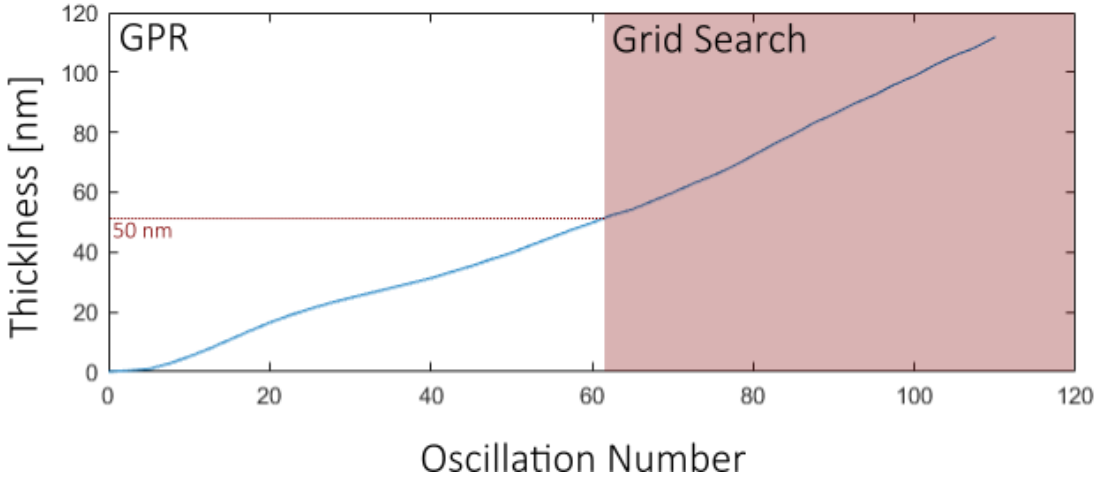

Figure 6.8: Fitted Thickness of ZnO film deposited at 150ºC vs Oscillation Number using the second-generation GPR algorithm.

algorithm doesn't need to extrapolate when used to estimate the thickness of thicker films. Therefore, as the estimates from the GPR and grid search algorithms were shown to reliably be equivalent after 50nm, this was the thickness selected as the transition point between the two search methods. While the GPR algorithm was still trained using the data from the grid-search algorithm, it is biased towards creating a smooth line with no "jumps", and tries to find the most likely smooth line based on what data it does have, the end result of which can be seen in Figure 6.8. In this example for a ZnO film deposited at 150℃, we can see that the initial spikes are gone, replaced with a smooth curve where growth rate steadily increases to the steady-state. In addition, instead of a nonlinearity as we leave the nucleation region, we now have a smooth curve showing growth starting at about 5 oscillations, and it continues to grow at around 1nm/oscillation as predicted, and there is a smooth transition at 50nm when switching from GPR back to grid-search. This result manages to capture the full growth mechanism of the film, by combining the two algorithms into one structure. Since grid-search methods search the whole data-space, they take a few seconds to find the best result, especially when not bound effectively. However, the GPR model was able to provide a reasonable thickness estimate in less time than it

takes for the reflectance spectrometer to take another measurement, on the order of a couple of milliseconds.

6.3.4 Automatic Feedback for AP-SALD System

By using GPR and GS algorithms in tandem, not only are our thickness estimates substantially more accurate, but it means that the film can be analysed in real-time. Conventional ALD processes can take several minutes to complete a deposition cycle, leaving plenty of time for a measurement to be taken. On the other hand, AP-SALD/CVD takes a couple seconds or less to complete each deposition cycle – necessitating this focus on measurement speed. Figure 6.9 shows how we were able to see the instantaneous response of an AP-CVD deposition performed at 150ºC to changing proportions of diethylzinc delivered to the reactor, something that could not be done if each measurement took longer than a single oscillation of the stage. In addition, because of this better fitting we can directly observe the nucleation period, deposit very

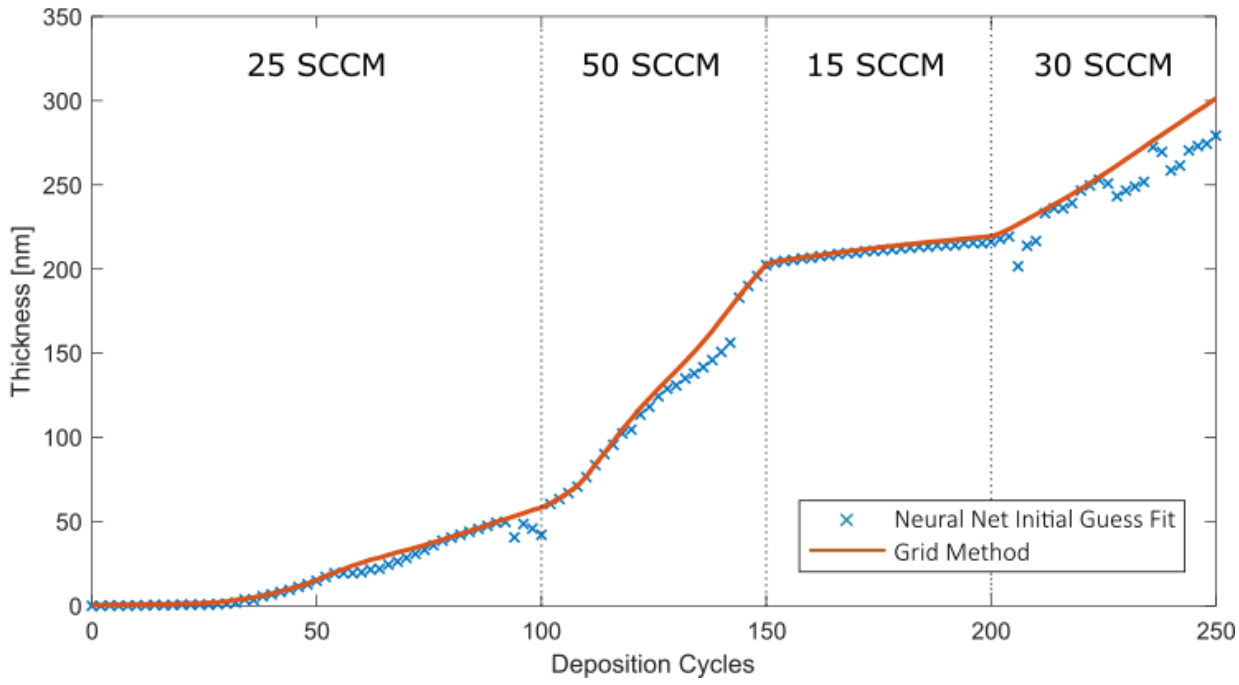

Figure 6.9: Changes in ZnO growth rate due to increased DEZ-laden nitrogen flow ratio through the bubbler, ratio of DEZ:Carrier indicated for each section – image courtesy of Kissan Mistry

thin films with nanoscale control (including nanolaminate structures), and very importantly, implement growth rate control and feedback mechanisms.

Because we now have accurate measurements of the film thickness through all stages of the deposition, it became feasible to implement a control system for the AP-SALD using this tool. We implemented a PI controller that would measure the thickness of the film at intervals of 5 oscillations, giving the growth rate time to adjust to alterations to the control variable. Since altering the temperature would alter the properties of the film itself, in potentially undesirable ways, we selected the flow rate through the DEZ bubbler to be the control variable. This works in the AP-CVD regime, but under ALD conditions the maximum growth rate is limited by the rate at which each layer is saturated. It is possible to lower the growth rate of an AP-SALD film by depositing less than a full layer each oscillation, but there is no advantage in doing so. We also know that once the growth rate has reached its new equilibrium, there are no processes that could cause a sudden change in the rate. Viewing the growth rate as the target variable and precursor bubbler flow rate as the control variable (*CV(t)*), it was possible to develop a control algorithm using only proportional and integral (PI) control, given by the formula;

$$CV(t) = K_P e(t) + K_i \int_0^t e(\tau) d\tau \tag{21}$$

In this case K_p and K_i are the proportional and integral gains, respectively, e is the error of the target variable at either time t, or at a point in the integration τ, between 0 and the current time t. The proportional control is the simplest part, but also provides the bulk of the adjustments to the control variable. Proportional control functions by simply taking the difference between the actual and desired values of the target variable, and multiplying that by some value (K_p) to get the amount by which the control variable (*CV*) should be changed. If K_p is too low, the error is

corrected too slowly, and if it is too high then the target variable will oscillate around the setpoint as the error value changes sign. Proportional control is also susceptible to what is called steady-state error, because eventually the error gets so small that the constant gain is not large enough to elicit a change in *CV*, while still being non-zero. This error can be very small, but since we desire growth rate control in part to attain precise thickness control under varied deposition conditions, it can add up over time, and this time-based accumulation of error is why we needed to add the integral control part of the algorithm. For this part of the algorithm, the gain K_i tends to be much smaller than K_p, as the error term is the *cumulative* error from the beginning of the deposition until the current time. That means that the longer the deposition, and the bigger the effect of the steady-state error, the bigger the integral term gets in response. Over time this will correct the steady-state error, and bring the growth rate to the target value, but because K_i is so small its effects are only noticeable over longer time-scales than those on which the proportional control operates. The values of K_p and K_i were tuned using the Ziegler–Nichols method, a mathematical technique where we set K_i to zero, and increased K_p until the output (measured thickness) began to oscillate. The value of K_p where this occurs is called the ultimate gain (K_u), and the period of oscillation is denoted T_u. By the Ziegler-Nichols method, we then set the proportional and integral gains to be $K_p = 0.45\ K_u$ and $K_i = 0.54\ K_u/T_u$. Through this method, we could reliably reach the growth rate setpoint within 15-20 oscillations after it was set – even accounting for the natural growth rate lag mentioned at the beginning of this section as shown in Figure 6.10.

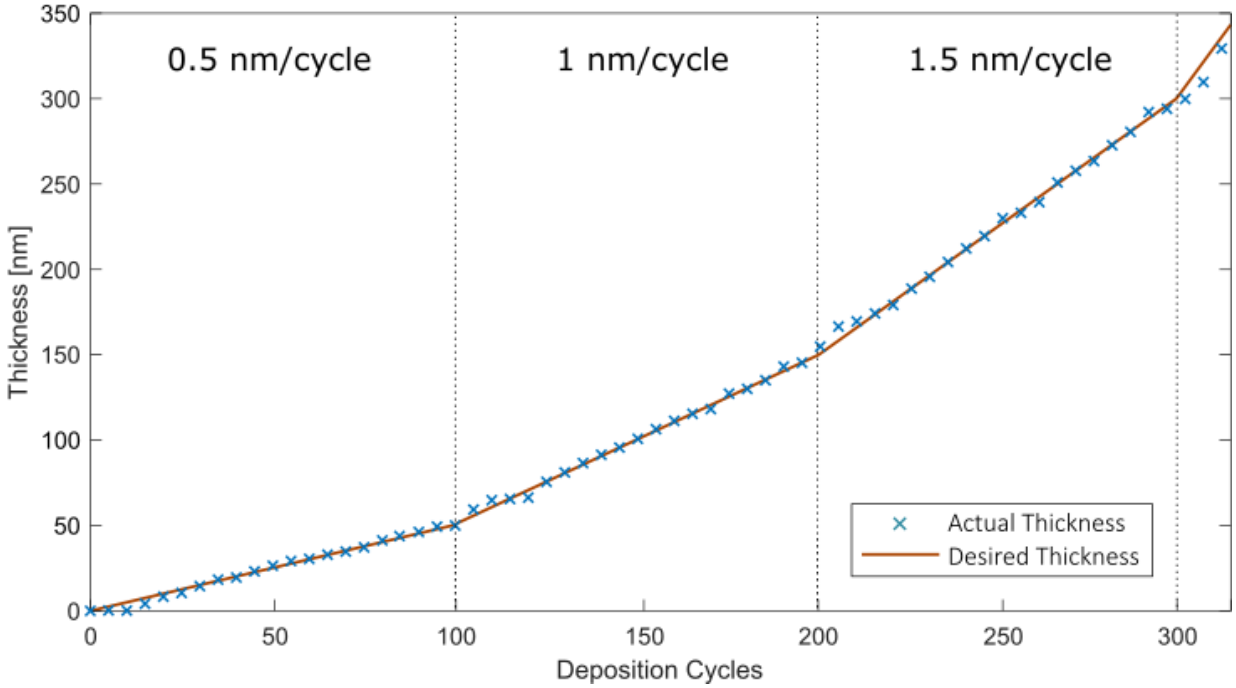

Figure 6.10: Response time of a ZnO film deposition at 150°C controlled by our in-situ PI algorithm

6.4 Conclusion

In conclusion, my collaborators and I tested several algorithms to allow us to use in-situ reflectance to accurately measure the thickness of ZnO films in real-time. By combining the power of a GPR model with the speed of a grid-search algorithm, and the accuracy LM fit we were able to reliably find the thickness of the film at all stages of the deposition. These measurements were fast enough that the deposition was not disturbed, and accurate enough that we were able to implement a PI controller which allowed us to much more precisely control the thickness of our films than was previously possible.

7 ML-Empowered Ellipsometry

7.1 Introduction

In contrast to the reflectance spectroscopy system presented in Chapter 6, the spectroscopic ellipsometry measurements herein were performed ex-situ. The FS-1 Ellipsometer acquired by the lab has the capability to measure and analyze across multiple sites on the film, a process that would take too long if performed in-situ. The onboard software also contains more advanced optical models, allowing us to extract a greater range of material properties, about a wider variety of films.

One of the primary advantages of our lab's AP-SALD system is its ability to deposit films with an astounding array of compositions. Multiple metallic precursors can be introduced to the reactor in varying proportions, allowing us to produce films that are just lightly doped, to degeneratively doped, all the way through multi-oxide films. However, the greater the complexity of the film itself, the more difficult it is to fit to an optical model through trial and error, as the number of parameters needed is increased and their ranges are less likely to be reported in literature.

Machine learning techniques have been applied to the problem of fitting optical models for ellipsometry data, however the main focus seems to be on simply determining the thickness [124], [203], [204] or refractive index [127] directly, bypassing the optical model. This is an excellent approach for industrial applications, where you know that the films are being produced with consistent properties, but it is only from the optical models that we could determine what those properties are if they are unknown. Because AP-SALD systems like ours have been used to deposit insulating [35], [205], semiconducting [31], [36], [206]–[208], and even metallic films

like the copper presented in Chapter 3, we must use an optical model to ensure that our films are being measured accurately. Various kinds of ML algorithms have been employed to attempt to solve the problem of fitting ellipsometry data to various thin films, and a common approach is to use classification algorithms like K-nearest-neighbours, logistic regression, and support vector machines [124], [204]. Classification algorithms are robust against over-fitting, as the final output is placed into one of a specified number of "classes", based on a probability value determined by the algorithm. However, this means that there will always be some degree of uncertainty in the predictions, as this method relies on constraining continuous variables into a discrete space. Deep neural network architectures have also been used to solve this problem without artificially discretizing the final answer [127], [203], but in order to obtain the necessary training data the researchers had to use simulated data. Since the intent of this study is to determine the relation between the AP-SALD process parameters and the composition and properties of our films, there is not yet sufficient understanding of the processes involved to create a suitable simulated dataset.

So far, the intended application of ML techniques to spectroscopic ellipsometry measurements has been fairly limited, and so it was appropriate to use the methods previously outlined. However, we wish to use ML to enhance our materials-discovery pipeline, and so we need a method that is both broadly applicable and designed to retain the expert in the loop, so confidence in its output may be maintained. The single longest stage when fitting a complex optical model by hand is finding initial conditions that are close enough to the film's actual properties that a conventional fitting algorithm can find the optimal values, and so we intend to automate this first step – accelerating materials discovery and retaining the confidence of applied expert knowledge.

The first stage of this project was to collect enough measurements of real films to begin training an ML model, and to that end metal-oxide films containing varying concentrations of aluminium, copper, zinc, nitrogen, and oxygen were deposited using our AP-SALD system onto 7x7cm borosilicate glass slides. Temperatures, flow rates, and the number of oscillations were all varied, and we leveraged our labs capacity to create thickness gradients [35] to maximize the information contained within each film – seeing how optical properties vary with thickness when composition is held constant with only a single deposition. Because the data was so varied, it was necessary to leverage transfer learning techniques to make sure that all input data was as accurate as possible. As previously mentioned, it is hard to fit a multicomponent film like nitrogen and copper-doped zinc oxide, but if we know the properties of a comparable ZnO film we can leverage that knowledge to obtain initial estimates for the film's properties. Different ML models are evaluated and a GPR model is found to work best. By combining a GPR ML model with a Tauc-Lorentz oscillator optical model we were able to get a baseline ML model, to which we iteratively added other optical models (Drude, Multi-Oscillator) until the full dataset could be analyzed.

Using the developed GPR model, we were able to predict the optical model parameters of a film based on its AP-SALD deposition parameters and raw ellipsometry data. The predicted optical model parameters were then used as initial guesses to accurately fit the ellipsometry data to the optical model using the ellipsometer's onboard software. This approach resulted in a 43.25% reduction in the average final fit difference compared to manual fitting and the processing time for samples of unknown composition was also significantly shortened, from hours to only minutes. The system was built with expansion in mind, and so by applying the

same transfer learning techniques already applied the model can be expanded to include materials not yet made in our lab.

7.2 Experimental Methods and Materials

7.2.1 AP-SALD Depositions

Using our AP-SALD/CVD system we deposited films of varying thickness and composition, by altering the ratio of precursors, deposition temperature, and number of oscillations. Where the general mechanics of AP-SALD/CVD depositions are covered extensively in Chapter 2, we made specific decisions as to the range of the deposition parameters. We deposited films at temperatures ranging from 50-250ºC, using flow rates of between 0-50 SCCM for diethylzinc and trimethylaluminium, 0-200 SCCM for water and ammonium, and 450 SCCM for Cupraselect™. The spacing between the reactor and the substrate was constrained to between 0.04mm and 0.15mm, with a difference of between 0-0.1mm between the two sides. As has been previously mentioned, one of the critical advantages of our AP-SALD system is our ability to seamlessly switch between ALD and CVD deposition modes by altering how much the precursor compounds are allowed to mix before adsorbing to the substrate. Figure 7.1a and b illustrate the difference between these two modes, with uniform AP-SALD films having lower growth rates due to the decreased contact between precursors; any additional precursor delivered to the substrate after a fully adsorbed layer is formed is wasted in AP-SALD mode. As the precursors are allowed to mix before adsorption in AP-CVD mode, the growth rate of the film is much more sensitive to the concentration and flow rate of each precursor. Because of this dependency on flow rate, we can deliberately alter the growth rate along the width of the substrate by altering how much the precursors mix at each point. This can be accomplished through the use of a specially designed reactor, or by simply tilting the reactor head as shown in Figure 7.1c, the

method we chose here. Uniform films can be produced using AP-CVD mode as well, but AP-SALD can *only* produce uniform films, and so the reactor is kept parallel to the substrate as shown in Figure 7.1d and f.

7.2.2 Sample Preparation

Because thickness has a very large effect on how a film responds to light when compared to other optical model parameters for films of similar composition, it is very worthwhile to make sure that the data for this variable is as comprehensive as possible. These gradient depositions allow us to create films with the same or similar properties across their width, while varying the thickness by over 100nm, with only a single deposition. To encompass all of this variability, each borosilicate glass slide was cut into 1x1cm squares, and each square was then measured in 25 locations by the ellipsometer as shown in Figure 7.1g. By taking so many measurements from each sample, we were able to leverage a feature of our FS-1 ellipsometer whereby a model is fit over all the points measured on a single sample. In this manner we can hold the value of some optical model parameters constant between measurements, while allowing others to vary. While this would theoretically decrease how well each individual point is fit if the optical properties of the film vary wildly, this technique can be very useful for finding accurate values of parameters that are harder to fit well. For example, the plasma frequency of degenerate AZO films tends to shoot to zero or infinity if measured over a single point, but can be more reliably found if analysed over multiple error surfaces. A film which has been fit using this multi-sampling method can be further refined measurement-by-measurement if that is necessary to improve the fitting, while retaining the values of those parameters that needed to be found using multiple measurements.

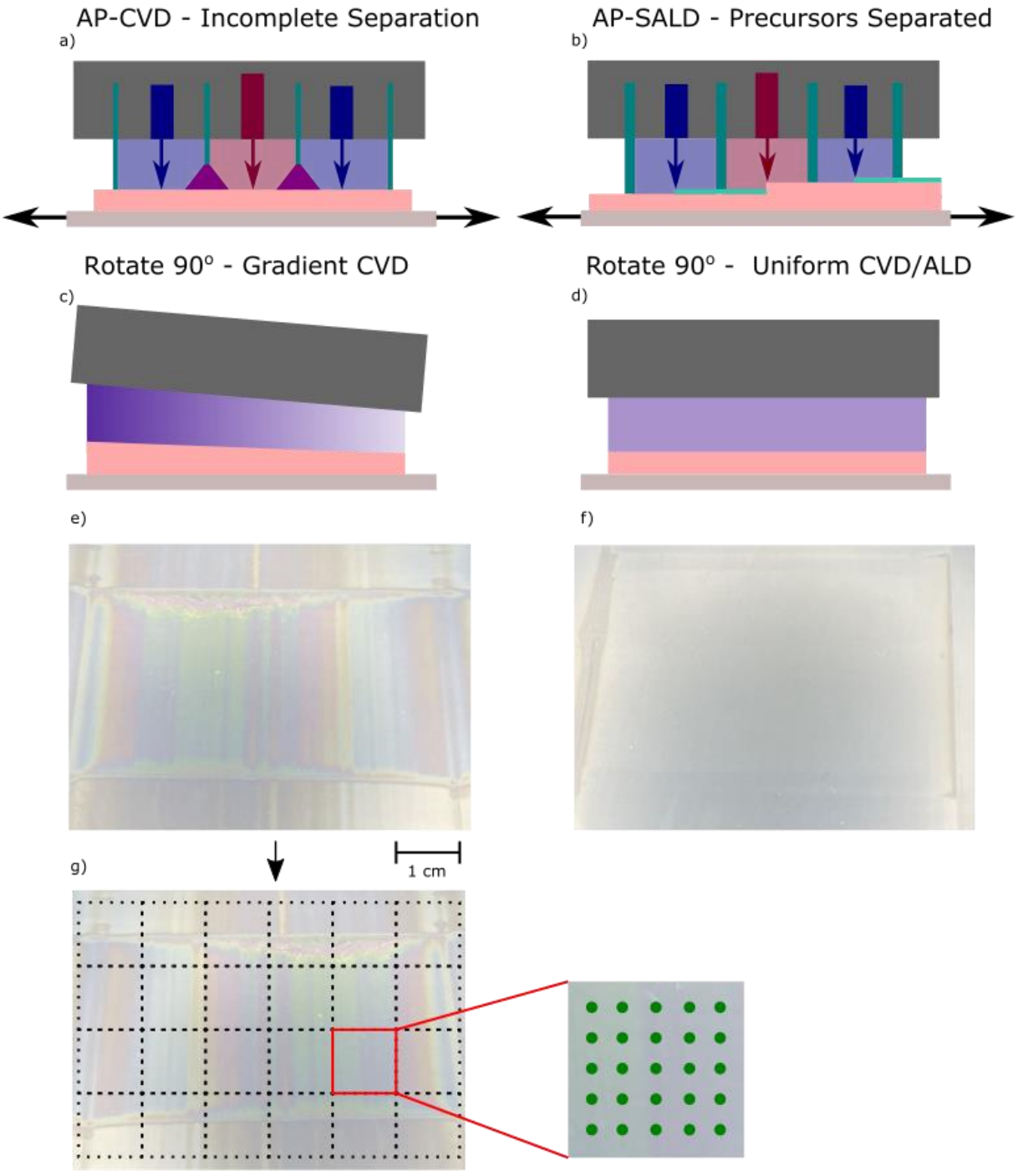

Figure 7.1: Schematic of how samples are prepared for inclusion within the training dataset. a) AP-CVD-mode deposition, with limited mixing occurring below the reactor head. b) A similar deposition occurring in AP-SALD-mode, where each precursor adsorbs fully before reacting with the next. c) How we can leverage the precursor mixing of an AP-CVD deposition to produce gradient films, where differing mixing rates lead to differing deposition rates along the width of the substrate. d) AP-SALD cannot produce gradient films, as the rate of reaction is limited by the complete adsorption of precursor onto the substrate. e) and f) AP-CVD gradient and AP-SALD uniform film produced at the same temperature of 200°C using 250 oscillations. g) Schematic of how each film sample is broken down and measured to provide hundreds of measurements per film.

7.2.3 Data Preparation

At each of these 25 points, the FS-1 provides a total of twelve ellipsometric parameters measured over six wavelengths, $\Psi_{1\text{-}6}$ and $\Delta_{1\text{-}6}$ as described in Section 5.1.2. Before any ML models could be trained, it was necessary to manually fit 3000 such measurements using the onboard program, to provide the initial training data. We started with zinc or aluminium mono-oxide films, as well as lightly-doped AZO films, as we have the greatest experience with fitting models for such samples given that they only require a Tauc-Lorentz oscillator to be accurately modelled. Any unused parameters were assumed to be zero in this initial training set, as adding unnecessary parameters can decrease fitting accuracy while increasing fitting complexity. When it was time to train an ML model, the raw ellipsometry data was combined with a vector containing the number of oscillations over which the film was deposited (N), the deposition temperature (T), and a set of binary values indicating which precursors were used to create the film ($P_{1\text{-}4}$). The selection of these additional parameters will be discussed shortly, but they each encode important information about the film being measured. Figure 7.2 shows how we combine these two sets of data into an input vector, which is then passed to our ML model, which gives an output consisting of the various parameters of the multi-oscillator model.

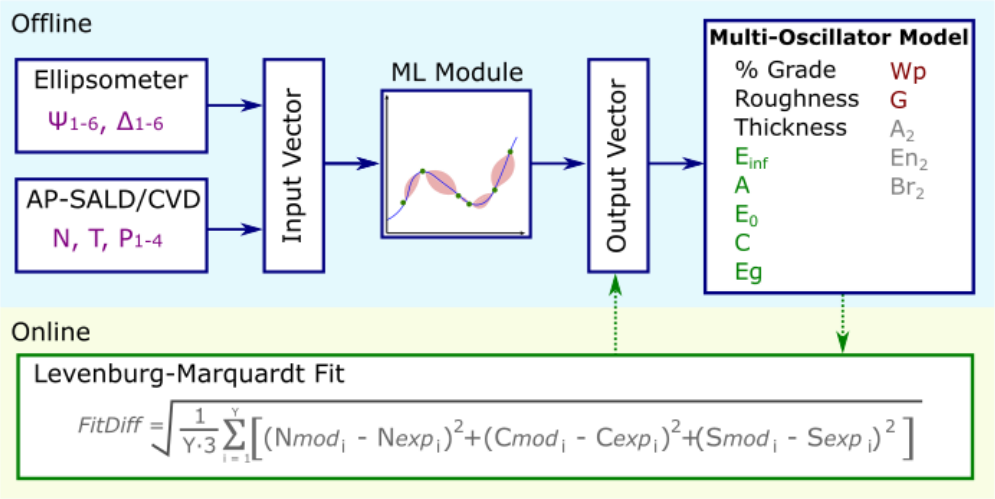

Figure 7.2: ML model workflow, showing how the output of the model is iteratively improved using the ellipsometers inbuilt LM fitting before it's used to train the next iteration.

In Figure 7.2, the various components of the multi-oscillator model have been differentiated by colour. In black are the universal parameters, which can be used in any model; thickness, roughness, and % grade. The first two are self-explanatory, but % grade is a mathematical abstraction that is very useful for AP-SALD/CVD films and will be explained in detail later. In green are shown the parameters of the Tauc-Lorentz model, which has already been discussed in Section 5.1.3, but in brief is a common optical model for semiconductors providing important information such as the material's band gap. In red are the two additional parameters needed to complete the Drude component of the multi-oscillator model, plasma frequency (w_p) and damping (G). These parameters are used to represent the behaviour of free electrons in metals and degenerate semiconductors, those which have been so highly doped that they exhibit metallic properties. Lastly, in grey, are the parameters for a Lorentz oscillator. Unused in this experiment, it is useful for materials which exhibit a second notable interband transmission peak greater than the band gap. So far this has not been shown to be necessary to accurately model any films we have produced, but it may be necessary for future materials added to our library.

7.2.4 Machine Learning Methods

In total, 13,326 ellipsometry measurements were taken to form the database that we used to develop our ML model, a figure not evenly divisible by 25 due to difficulty physically breaking the films into perfectly square pieces. Of this total, 3000 were fit manually and from these 3000 measurements, 1000 were selected to compare various ML models, and another 1000 measurements were selected to determine what additional data from the AP-SALD would improve the models' performance (these sets of 1000 were allowed to overlap). This distribution of initial training samples is shown in the upper portion of Figure 7.3. First, we compared the RMSE and R^2 values of four different ML models trained using thirty cycles of Bayesian

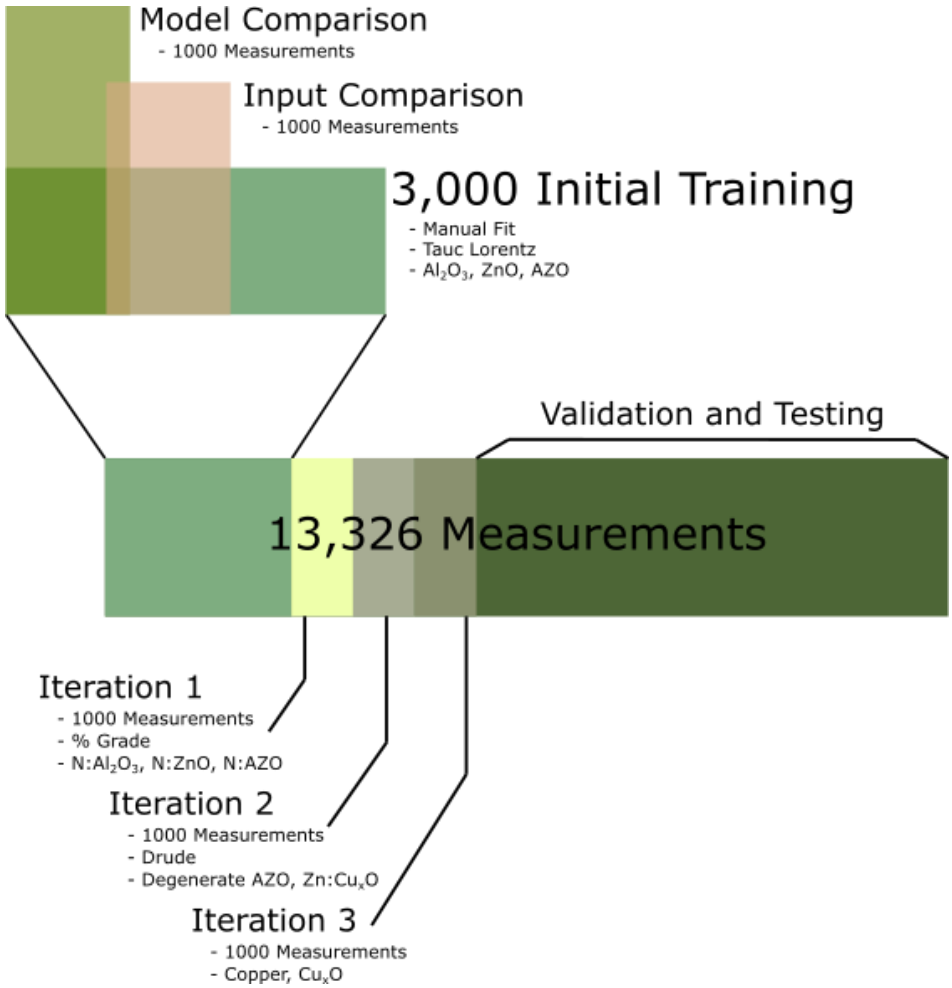

Figure 7.3: Division of the total measurements taken for this work, the bulk were reserved for validation and testing, but the differences between various iterations of the model are shown.

hyperparameter optimization; linear regression, random forest, artificial neural network, and gaussian process regression. The first three model types have been used by prior researchers to solve similar problems in ellipsometry, and based on prior experience we know that GPR can perform very well in these tasks. Comparing the various models in this initial stage will help us determine if we are moving forward with the best technique for solving our problem, and training is comparatively fast because the dataset is small. In addition, because the dataset is small, models that are prone to overfitting will do so, and the models which are better at predicting the underlying relationships instead will be more apparent. After we found the model that best represented the data, we explored whether the AP-SCVD input data we were adding to the raw ellipsometry inputs was affecting the accuracy in any meaningful way. Using the model selected in part 1, we trained a version of it with and without all the additional input data (N, T, P1-4). Minimizing the number of inputs broadens the applicability of the model outside our lab – the relationship between say, temperature and roughness may be unique to our system, but each additional piece of data carries a lot of additional information about the film. However, the actual effect of each of these data on the accuracy of the final model needs to be determined. Once we understood which model worked best, and that our inputs were optimal, we proceeded to train an initial version of the ML model using all 3000 measurements. This version used the raw ellipsometry data, deposition temperature, number of oscillations, and composition as inputs. After training it using five-fold cross validation and stochastic gradient descent we were able to get a model that would provide us with the Tauc-Lorentz model parameters, thickness, and roughness of the film.

As discussed in Section 5.2, transfer learning lets you expand upon an existing model by training it on new data, using the existing weights and biases within the network as an initial

guess. This works best the more closely related the two sets of training data are, and ours are very closely related because the films all have certain elements in common. There is always going to be a high correlation between the number of oscillations and thickness, no matter how else the film changes, for example. Each iteration was trained using data from the previous versions, as well as 1000 new samples that highlight the new properties we are looking to predict. We chose to expand the model in the order that we did, to leverage our experience with fitting ZnO and alumina films using the Tauc-Lorentz model. Because we are using transfer learning to iteratively expand the model's capabilities, we need to make as sure as possible that the ML model is being trained on accurate data. Because we know the ranges of all the Tauc-Lorentz parameters that we expect to see in these films, we can confirm that the model is working as intended. We then expanded the model to include the % Grade (Iteration 1 in Fig Figure 7.3), as we knew that this was a parameter that would be present in all our films by the nature of AP-SALD, and it has a large impact on the fit difference. After that we expanded the model outputs to give us the Drude parameters, plasma frequency and damping (Iteration 2 in Figure 7.3), by training the model on degenerate semiconductors and Cu-doped ZnO. The composition of these films is very similar to those of the prior iterations, but their properties are such that they do have free electrons. Lastly, as we had introduced films with some copper content in the prior iteration, we expanded the model inputs to include films that were primarily or entirely copper-based (Iteration 3 in Figure 7.3). These films have a much stronger dependence on the Drude parameters than the Tauc-Lorentz, unlike the other films. To ensure that the full range of copper compositions was represented in the data we deposited pure Cu films of known thickness in a thermal evaporator, then measured them using the ellipsometer. We then oxidized them and repeated the process several times, allowing us to fit an optical

model and focus only on the unknown parameters, such as the plasma frequency and the band gap of the oxides. Some copper oxide films were deposited using the AP-SALD and included in this training set, and we used our experience fitting the evaporated copper to ensure model validity. For both manual fitting and when verifying the ML model, it is necessary to make sure that the parameters are within reasonable bounds, as confirmed in the literature and presented in TABLE V and TABLE VI.

TABLE V: BAND GAP AND INTERBAND TRANSMISSION VALUES FOR MATERIALS USED IN THIS STUDY

Material	Eg (eV)			
ZnO	3.18-3.29	[209]	4.105	[50]
	3.13	[210]	3.4	[214]
	3.27	[21]	3.15	[215]
	3.37	[17], [211]	3.28	[216]
	3.3	[212]	3.51	[217]
	3.31-3.35	[213]	3.35-3.39	[218]
Al_2O_3	5.5	[219]	6.5-7	[222]
	7.3	[220]	4.3	[223]
	3.88	[221]	4.64-6.24	[224]
AZO	3.4	[52]	3.06-3.12	[215]
	3.65-3.76	[225]	3.3-3.48	[216]
	3.08-3.325	[209]	3.62-3.69	[228]
	2.5	[226]	3.87	[217]
	3.25-3.29	[227]		
Cu	2.6, 3.9, 5.2	[229]		
CuO	1.986-2.355	[230]	1.2-2.7	[232]
	1.2-3.85	[231]	2.03-2.8	[233]
Cu_2O	1.92-2.02	[234]	2.7-3.02	[233]
	2.1	[235]	1.93-2.48	[237]
	2.58	[236]	1.2-4.9	[238]
	2.0	[52], [232]		
Cu:ZnO	3.444-3.954	[50]	1.5-2.5	[239]
	2.64-3.76	[209]		

As can be seen, the same material can have very different properties depending on morphology, stoichiometric ratio, and deposition technique. Alumina, for example, has reported band gap values ranging from 3.88-7.3 eV which is because defects in the material can provide energy levels within the band gap of the bulk, allowing electrons to move to the conduction band. All of the films in TABLE V and TABLE VI were produced using thin-film deposition methods, but there is not enough data on films produced using AP-SALD/CVD to rely on that alone. Different deposition techniques are susceptible to different kinds of defects, as we saw in Chapter 0 when the atmospheric nature of our AP-SALD/CVD induced a greater than expected number of oxygen interstitials in the crystal lattice of the ZnO. Understanding the intricacies of how these material properties are found, and what can affect them, is a large part of why retaining an expert in the loop is so critical for this application.

TABLE VI: DRUDE PARAMETERS, PLASMA FREQUENCY AND DAMPING FOR ALL RELEVANT MATERIALS

Material	AZO		Cu		ZnO		CuO	
Wp (eV)	0.6529-0.6563	[225]	7.98-9.30	[243]	0.0728-0.3658	[246]	4-5	[247]
	0.892	[240]	7.39	[244]				
	1.7473	[241]	8.76	[245]				
	0.9537	[226]						
	0.35-0.7	[40]						
	1.28-1.7	[242]						
	7.61-9.59	[228]						
G (eV)	0.04486-0.064	[241]	0.0092-0.126	[243]	N/A		0.01-0.5	[247]
	0.01-0.33	[40]	0.00446	[244]				
	0.096-0.28	[242]	0.0477	[245]				

The ML model is blind to what an acceptable band gap for ZnO is, and some films that we analysed have much higher or lower band gaps. If the ML model provides an estimate of the band gap of ZnO that is far outside the bounds presented in TABLE V, then I know that the result is an error that will have to corrected during the manual refinement phase. If this happens, we pin the value of the erroneous parameter to something that is reasonable based on our knowledge of our films, say 3.3 eV for the band gap of ZnO. Then we fit all of the other model parameters without allowing the pinned value to change, iterating through them all until the fit difference is reasonable and the other parameters are also within a reasonable range. We fit the erroneous value after all the other parameters because the error surface of the Levenberg-Marquardt algorithm changes as all the parameters change. We assume that the pinned value is close to the error minimum of this surface, but not at the minimum itself and so allowing the other parameters to adjust first means that the Levenberg-Marquardt algorithm is more likely to find the actual minimum.

7.3 Results and Discussion

As previously mentioned, the first stage of developing the best possible ML model for our application was testing various kinds of ML models, and comparing their performance. Of the four models we compared, three had been used by prior researchers to solve similar problems in ellipsometry measurements, although none had tried to create a system that produced all the parameters of an optical model. The last was a GPR method, with which we have experience working, and seems well-suited to this kind of problem. After all, the optical model that we are fitting is a smooth mathematical equation describing the behaviour of the material, based on its optical properties, and all our data falls close to the true value of the equation because it is constrained by the acceptable fit difference. As part of the hyperparameter optimization process,

Matlab would automatically test the performance of four different GPR kernels in order to find the covariance matrix; rational quadratic, squared exponential, Mattern 5/2, and Exponential. TABLE VII summarizes the results of this stage of the experiment, and it can be seen that overall the GPR model performed the best with the exception of the % grade parameter. It was at this point that the decision was made to include % grade in a future iteration of the ML model, because much as with the reflectance data in Chapter 6, the better-performing ANN was likely learning the noise present in the original data. Here the noise came from the fact that % grade was not needed to get an acceptable fit for many of the films in this initial data-set, and attempting to add it unnecessarily would increase the complexity of the fitting and add potential errors. Leaving it out of those measurements where it was unnecessary was the best decision for preserving the integrity of our dataset.

TABLE VII: R^2 AND RMSE VALUES OF FOUR ML MODELS COMPARED, BEST RESULTS HIGHLIGHTED IN BOLD

	Linear Regression		Random Forest		Artificial Neural Network		Gaussian Process Regression	
	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
% Grade	0.50	5.9636	0.54	5.739	**0.94**	**2.087**	0.85	3.257
Roughness (nm)	0.93	2.617	0.87	3.709	0.94	2.613	**0.97**	**1.731**
Thickness (nm)	0.97	25.129	0.96	26.35	0.98	19.19	**0.98**	**18.51**
E_{inf} (eV)	0.11	1.077	0.59	0.734	0.76	0.565	**0.81**	**0.498**
A	0.60	40.99	0.79	26.67	0.86	23.94	**0.90**	**20.52**
E_0 (eV)	0.60	1.731	0.62	1.675	0.85	1.067	**0.85**	**1.038**
C	0.85	2.313	0.82	2.545	0.82	2.551	**0.93**	**1.601**
Eg (eV)	0.88	0.1597	0.93	0.1194	0.96	0.089	**0.97**	**0.077**

Linear regression performed worst across the board, which makes a lot of sense in this case. While we have a regression problem, none of the output variables are linear with respect to the inputs, with the exception of thickness being linearly dependent upon the number of cycles. GPR makes statistical assumptions about the spread of the data points from some real value, and does not constrain the shape of the relationship. The ANN was also performing linear regression, but

because there are multiple layers of linear regression the individual linear functions were able to sum to a nonlinear relation, resulting in improved performance.

At this juncture, it should be explained what exactly % grade is, and why it matters to our films. It is a mathematical representation of how the optical properties of a film can change between the surface and the bulk, and AP-SALD films are expected to have this sort of graded profile for a couple of reasons, illustrated in Figure 7.4. Firstly, incompletely reacted precursor ligands at the surface; these would normally desorb during later deposition cycles, but this can leave some at the surface after a deposition is complete. There are also usually unreacted hydroxyl groups at the film surface, producing a thin reactive layer. Secondly, AP-SALD/CVD depositions require elevated temperatures in atmosphere, representing a chance for oxygen interstitials and vacancies to form and migrate. As the initial layers are annealed for longer and sealed from the atmosphere by subsequent layers, these processes have more time in which to occur. This effect is expected to be more prominent in thicker films.

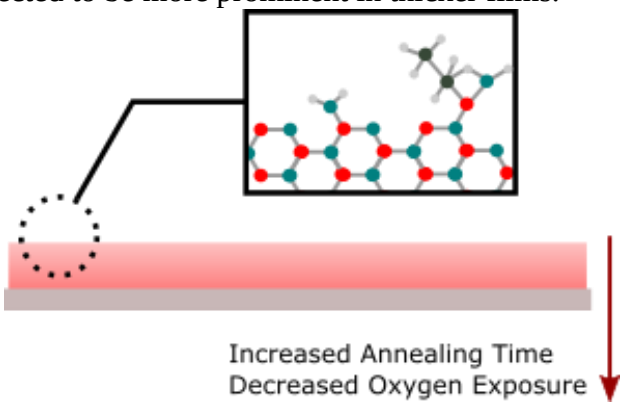

Figure 7.4: Illustration of the common causes of graded optical properties in AP-SALD films.

This is a large part of what makes AP-SALD films unique, but also uniquely challenging to fit with conventional methods. In particular, any copper-containing film will have a graded optical profile, as the surface oxidizes very quickly, but all our films are annealed in oxygen as an inherent part of the deposition process. This parameter also must be fit last, as it can have a very

large effect on a film's fit difference and if you find the % grade first you can create an artificial error minimum, and not represent the true properties of the film.

Now that we showed that our initial assumptions about the suitability of the GPR method to solving our problem were valid, it was time to determine if the data we collected from the deposition process itself was improving the model or making it needlessly more complex. We trained a GPR model five times and averaged the validation RMSE of each trial in an attempt to make sure the differences between the values were not a result of a single poorly initialized training session. We did this first using only the raw ellipsometry data, then iteratively added deposition parameters to the training. TABLE VIII compares these average RMSE values over three of the most important model parameters, thickness, roughness, and band gap – as these not only have a large effect on the model performance but are also properties we are actively interested in determining. We already knew that, of the known deposition parameters, composition, number of cycles, and deposition temperature have the greatest effect on the material properties of the film. For example, we expect that a film which has been deposited for 1000 cycles will be significantly thicker than one deposited for only 500, but that the degree of this difference will also depend on the temperature and material being deposited. We do not need to know the relative strength of these interactions, only whether or not the model has been improved by their inclusion.

TABLE VIII: RMSE COMPARISON OF GPR MODELS PRODUCED WITH COMBINATIONS OF DATA ABOUT THE DEPOSITION PROCESS, BEST RESULTS SHOWN IN BOLD.

	Raw Only	+ Composition	+ Temperature	+ Cycles	+ Composition + Cycles	+ Composition + Temperature	+ Temperature + Cycles	All
Roughness (nm)	3.325	3.227	3.033	3.139	3.159	3.080	3.166	**2.147**
Thickness (nm)	15.58	14.70	14.24	14.01	13.92	14.04	14.03	**13.83**
Bandgap (eV)	0.1130	0.1118	0.0693	0.0746	0.0684	0.0655	0.0645	**0.0625**

We can see from this table that, for example, adding the number of cycles alone improved the RMSE of the thickness and adding the deposition temperature almost halved the RMSE of the predicted band gap. However, the version that included all three deposition parameters had a slight advantage over any other combination, which again makes sense given that all three have a significant effect on a film's properties.

Figure 7.5: Average fit difference of all films in each iteration

Now that we are confident the model has been initialized as well as possible, it was time to begin the process of iteratively expanding the capabilities of the model. Figure 7.5 below shows the average fit difference of each series of measurements *after* manual refinement, as the inbuilt ellipsometry software was not able to provide pre-refinement fit differences for comparison. The fit difference of the manual iteration and three transfer learning iterations is based on the training data for each, and the testing fit difference is based on a random selection of 1000 measurements that were not used to train any part of the model. We can see that the largest reduction in the fit difference comes with the addition of the % grade parameter, but as discussed previously this is expected as it is a component of all our films produced using the lab AP-SALD system.

There is a slight rise upon introducing the Drude parameters, but the fit difference is still excellent for films of this thickness. It is most likely due to the fact that the Drude parameters had a strong tendency to either collapse to zero or explode to their set maximum value, making them slightly more difficult to fit. This most often occurs in the degeneratively-doped semiconductor films, so is likely due to the comparatively small number of free electrons compared to metallic films. The second, slightly larger rise for the testing dataset is just because

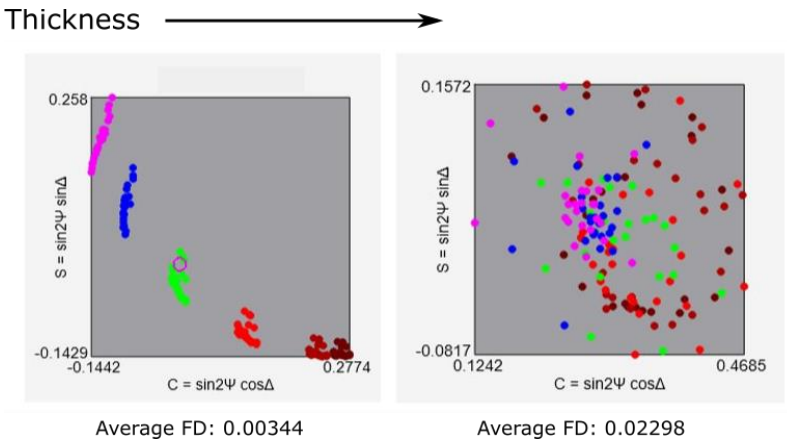

Figure 7.6: Raw ellipsometry data and average fit difference of two films used in this study. On the left is a 50nm ZnO film produced under uniform ALD conditions, and on the right is a Cu:N:ZnO film produced under gradient CVD conditions with a thickness ranging from ~250-300nm.

none of the samples contained therein were used for training at all. It is still an excellent result, and the fact that the loss of accuracy is so low for such a comparatively small dataset shows how representative we were able to make the initial stages of the training data. Comparing the fit difference of the testing and manual sets, we can see that the average fit difference has been reduced to only 43.25% of the original value and that with a much broader range of film compositions compared to the manual data.

In addition to improved accuracy, by using this algorithm, we were able to reduce the time to fit of a four-precursor film from over an hour of manual fitting, to less than ten minutes. This is because we are providing an initial guess to the ellipsometers Levenberg-Marquardt algorithm that is significantly closer to the actual values of the film, even if that film's exact composition is unknown. We can also expand this model to new precursor compounds as needed, using the same transfer learning technique. Figure 7.6 shows an example of how fitting different samples can vary in complexity. Given that each sample needs to be fit iteratively, there is a significant difference in the length of time it takes to fit the film shown in the image on the left vs. the one on the right, and even so there can be a loss of accuracy when manually fitting a more complex or thicker sample. The image on the left is a ZnO film, approximately 50nm thick, produced under AP-SALD conditions. Accordingly, it is very uniform, and the variation between measurement points can be accounted for by altering a single variable – in this case thickness. All other model parameters could be held constant, and we obtained an excellent fit difference across all 25 points of 0.00344. The image on the right, on the other hand is the raw data from a Cu:N:ZnO film produced under gradient AP-CVD conditions. There is no consistent pattern visible in the raw ellipsometry data, and not only did each point need to be fit individually – all of the MO model parameters had to be allowed to vary. That represents 10 times the parameters

that needed to be fit, over 25 times the number of iterations. While the properties of each measurement could be used as a baseline estimate for fitting the next, this still relies on each fitted parameter being accurate across all 25 samples. Being able to reduce the number of iterations needed to get an accurate estimate of the parameter values through the use of ML represents a time saving of more than an hour of work.

One of the most significant challenges facing the use of ML techniques is the so-called "Black Box" problem, where we cannot ascertain how a given model reaches the conclusions that it does. This has proven to be a problem in the past, especially when datasets used to train a model contain implicit biases not recognized by the programmers working on it even leading to the perpetuation of systemic issues such as sexism and racism [248]. Fortunately, our model is only working in materials science, and isn't likely to cause harm to anyone but we still need to ensure that the results it has given are reflective of the reality of our films. As previously discussed, even after Levenberg-Marquardt refinement, it is possible to find an error minimum that is not reflective of the real properties of the film. To check the accuracy of our optical model, we used both atomic force microscopy and absorbance measurements to test the roughness and band gap of our materials respectively.

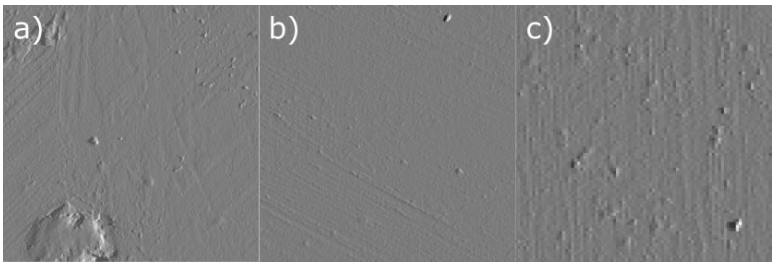

Figure 7.7: AFM image procured from three different 10x10μm areas of an AZO film. a) Edge of the deposition area, b) centre of the deposition area, and c) outside the deposition area, where precursors flowed out from below the reactor.

Figure 7.7 shows three AFM images prepared using Gwyddion, all taken from the same film in different locations, the film was deposited at 175ºC for 200 oscillations, with 23 SCCM DEZ, 10 SCCM TMA, 45 SCCM H_2O, and a total volumetric flow of 150 SCCM in each reactor

channel. There are multiple measures for roughness, none of which are used directly by the ellipsometer which approximates roughness using the Bruggeman Effective Medium Approximation (EMA) as discussed in Section 5.1.3. The two relevant for this section are the root mean square (RMS) roughness, measured against the mean line of the surface, and average peak height (APH), which considers the average valley depth. The formula for the APH roughness is self-explanatory, but if we define (l_r) as the measurement length and ($z(x)$) as the profile height function, the RMS roughness is given by:

$$R_{RMS} = \sqrt{\frac{1}{l_r}\int_0^{l_r} z(x)^2 dx} \tag{22}$$

Figure 7.7a was taken from the very edge of the deposition area, where we expect roughness to be higher due to oxygen ingress at the sides of the reactor causing reactions above the surface of the film, which can be verified by the presence of a large powder inclusion in the lower-left. This sample had an average roughness of 14 nm according to the ellipsometer, and when analysed using the AFM and Gwyddion the RMS roughness was 14.6 nm. Figure 7.7b is a sample from the centre of the film, and had an EMA roughness of 6.8 nm, which correlated well with the RMS roughness of 6.9 nm determined by AFM, while having a similar thickness to the sample from Figure 7.7a. Figure 7.7c was taken from outside the main deposition area, and so the film is formed of distinct islands, with areas of bare glass visible as deep striations between, a feature that is not visible on uniform ALD or CVD films. This sample had an EMA roughness of 24.7 nm, which was much higher than the AFM-measured RMS roughness of 8.1 nm, but aligns well with the average roughness peak height of 24.4 nm determined by AFM. This makes a lot of sense, because RMS roughness is calculated using the height deviations from the mean line, and areas with no film would by necessity bring the mean value down without representing much

about the film, whereas the average peak height is measured compared to the valley depth – in this case the substrate. While it was infeasible to perform these measurements for all our samples, there was very good agreement between the EMA and measured roughness for all the samples we did measure.

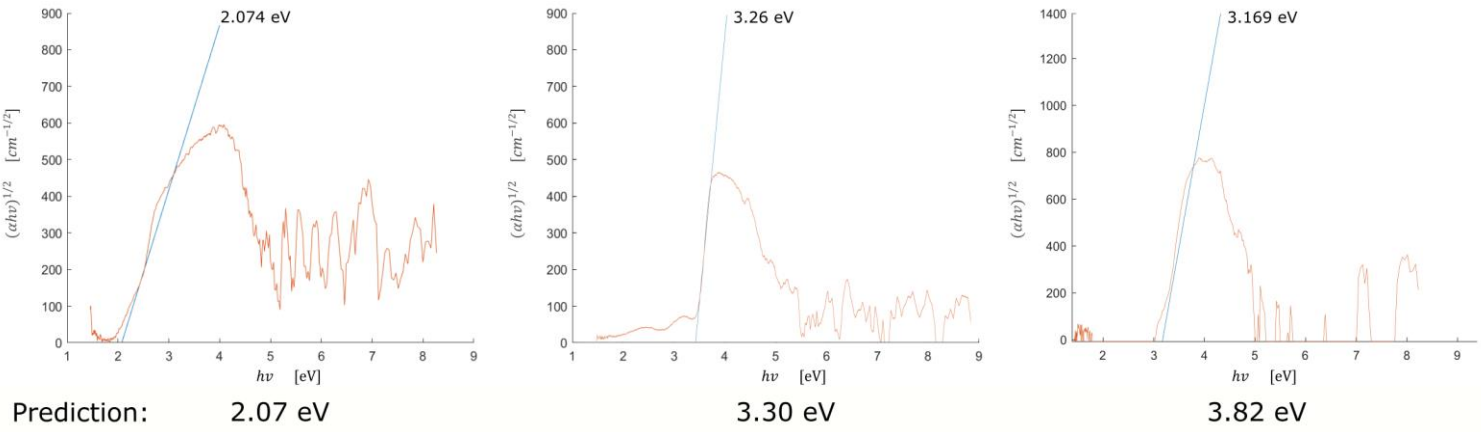

Figure 7.8: Tauc plots of three samples of varying composition and thickness, from left-to-right we have Cu_2O, ZnO and AZO films with the band gap indicated in blue.

We are also interested in confirming the optical band gaps of our materials, as this is a parameter of great interest when designing new materials for specific applications. By using absorbance spectroscopy and the film thicknesses provided by the ellipsometer, we can produce a Tauc plot giving us the optical band gap as shown in Figure 7.8. For a random selection of twenty samples from all iterations, the band gaps provided by the ML-assisted ellipsometry model had very good agreement with the measured band gaps, with a testing RMSE of 0.108 eV. This difference can be partially explained by the fact that it was only possible to take one absorbance measurement per sample, and so we had to compute the Tauc plot using the average thickness from the ellipsometer model, and compare it to the average band gap from the 25 ellipsometry measurements. The above examples are very good matches, selected from samples with comparatively uniform thicknesses. Alas due to that limit of the absorbance spectrometer,

films that had a large variation in thickness did not produce an absorbance band gap as close to the ellipsometer-provided value. On the left, we see one of our copper oxide films which had a measured band gap of 2.07 eV, and after refinement the ellipsometry model predicted an average band gap of 2.07 eV. In the centre we have a simple ZnO film for which the measured band gap of 3.26 eV is again very similar to the result provided by the ellipsometer model, 3.30 eV. However, on the right we see an example of a potential multi-oxide film, produced using our alumina and zinc precursors. The measured band gap is 3.17 eV, consistent with AZO in the tables above but closer to the range of pure ZnO than Al_2O_3. After refinement, the ellipsometer provided a band gap of 3.82 eV, still consistent with AZO – but outside the bounds of ZnO's commonly reported band gap values. These are both reasonable results for the bandgap of an AZO film of unknown composition, and there is no way to determine which is more accurate without further testing to determine the exact composition of the film, which would be prohibitively expensive to perform on such a large database. This highlights the challenges that we face when using optical techniques to measure a film's properties, and why retaining the expert in the loop will always be better than naively trusting any algorithm.

7.4 Conclusions

Through the course of this project I was able to develop a GPR model that allows us to analyse films of arbitrary composition with our ellipsometer with greater accuracy and speed than previously possible, which can be applied to other groups with similar deposition systems. The underlying assumptions of the GPR model mean that it is well-suited for working in concert with the LM algorithm used by the ellipsometer software to find the optical model parameters that best describe a film. In addition, we ensure that this procedure provides accurate results by retaining an expert-in-the-loop to verify the predictions of each parameter based on domain

knowledge. The average fit difference of our samples fell from 0.0326 to 0.0141, and the time to fit a sample fell from an average of over an hour, to less than ten minutes. This is because each parameter now has a more accurate initial estimate for the LM algorithm to refine, provided by the GPR model, rather than starting from scratch each time.

8 Computer Vision for Perovskite and Calcium Film Degradation Detection

8.1 Introduction

Many nanoscale devices contain compounds which are sensitive to exposure to air and water, rapidly degrading their performance. The commercial adoption of one such device structure is currently limited almost entirely by this degradation; perovskite materials can be used to construct photovoltaics (PV) that are not only incredibly efficient, with all-perovskite tandem PV's theorized to reach an efficiency close to or even exceeding 30% [249], [250], but they are also much cheaper to produce than silicon PV's. However, silicon PV's can remain in service for over 25 years before needing to be replaced, whereas first-generation perovskite PV's degraded in minutes [251]. Current developments allow for cells which retain their efficiency for 1000's of hours under ambient conditions [252], however, this still fails to reach the threshold for commercial viability, and often comes at the cost of the device's efficiency.

Encapsulation of environmentally-sensitive devices is a promising method for increasing their longevity, particularly films with a low water vapour transmission rate (WVTR) that can isolate the active layer from water for as long as possible. Conventional ALD has previously been used to produce thin film encapsulation layers that have very low water vapour transmission rate (WVTR) values [253], [254], some of which are also capable of withstanding mechanical stress with low loss of performance [255] due to the very high quality of the ALD film. Machine learning techniques have already been utilized to optimize and design other kinds of encapsulation layers, for example Hartono et. al. used a random forest algorithm to determine the ideal composition of such a layer [256]. Interestingly, they were able to parameterize non-

numerical data like the colour of the perovskite as a continuous numerical variable, allowing for the implementation of a regression algorithm. With most degradation mechanisms, colour change is a critical component of identifying when a desired compound has been replaced by one that is detrimental to device performance. In the case of perovskite films this can be achieved with absorbance measurements, but can also be detected with the naked eye. ALD films are ideal for this kind of encapsulation because they conform to the deposition surface near-perfectly and can be formed without pinholes and few defects, flaws which can make an otherwise well-isolating material more susceptible to the ingress of water. However, conventional ALD is an expensive and time-consuming deposition process, not only invalidating the commercial advantages of perovskite PV's, but making these encapsulation layers difficult to test, as holding a perovskite film at elevated temperatures for the length of a conventional ALD deposition may cause irreversible damage to that film before any testing can occur. Perovskite materials also often include highly-toxic lead, so it would be more sustainable to find an analogous material upon which to perform these experiments. In addition, the better a barrier performs the harder it is to get an accurate idea of exactly when it begins to fail. Early perovskite cells failed in minutes, but if the cell is able to last for days there will inevitably be gaps in coverage if manually imaging the progression of its degradation, particularly overnight.

First, I developed a spatially-resolved CNN which was able to successfully track the degradation of a perovskite film sealed with an AP-SALD encapsulation layer that had varied properties along one axis. This method was then improved upon by the introduction of calcium WVTR samples which could be tested more safely, and a new CNN was trained with greater spatial resolution to provide an accurate estimate of the area of the film which has degraded. By developing a method to track the degraded portions of a film in real-time, we can potentially

overcome a major limitation of calcium WVTR testing which is the assumption that the rate of degradation is linear. In the case of the perovskite films, the single greatest advantage this work provides is precise tracking of barrier effectiveness over several days. In combination with our gradient AP-CVD process, we can now more effectively design barrier layers, and obtain precise knowledge regarding their performance. In the case of the calcium WVTR testing, we have allowed for real-time, autonomous monitoring of the rate of change of the degraded area.

8.2 Experimental Methods and Materials

8.2.1 Sample Preparation

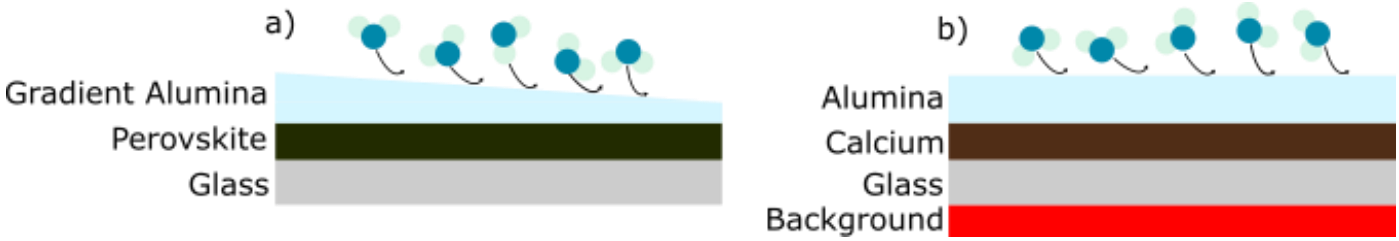

Figure 8.1: Film stacks used in phase one and two of this work, showing a) perovskite films with gradient alumina and b) calcium films with uniform alumina and added background layer.

Figure 8.1 illustrates the schematics of the two film stacks used in this work, including the gradient and uniform alumina barrier layers and the active layers of perovskite and calcium in a) and b) respectively. The perovskite layer is 315 nm thick, and is composed of formamidinium lead iodide, which was spin-coated onto a borosilicate glass slide. The perovskite layer was fabricated in a nitrogen glove box by combining:

- 1.4 M lead iodide
- 1.25 M formamidinium iodide
- 0.0987 M caesium iodide
- 0.0833 M guanidinium iodide
- 76% volume dimethylformamide

- 24% volume dimethyl sulfoxide

This solution was filtered to remove any solute or contaminant particles larger than 0.45 μm. At the end of the spin-coating process, the perovskite layer was precipitated using a 50/50 solution of chlorobenzene and chloroform, and then annealed at 140ºC for half an hour. This layer was coated with a gradient-thickness zinc-doped alumina film using our AP-CVD in the same configuration as for the films used in Chapter 7. This film was deposited at 130º to prevent damage to the perovskite layer, and was produced using 38 SCCM of TMA, 10 SCCM of DEZ, combined in the same channel with a nitrogen carrier flow of 235 SCCM. Water was used as the oxidant, which was introduced at 100 SCCM, combined with 275 SCCM of carrier nitrogen. The gradient thickness, ranging from 40 – 70 nm was produced by varying the reactor height on one side from 100 – 200 μm during a deposition of 60 oscillations. The 100 nm calcium layer was deposited in our evaporator, and topped with a similar alumina layer to the perovskite samples above but in the uniform AP-SALD configuration, these alumina layers had varied thicknesses, being deposited for between 50 – 300 oscillations.

8.2.2 Film Degradation Detection

The perovskite film, shown below in Figure 8.2, degrades through the formation of lead iodide, catalysed by the presence of water within the film. The visual change as this reaction occurs in the film is noticeable with the naked eye; the film changing from green-black to yellow as the lead iodide precipitates out of the perovskite layer.

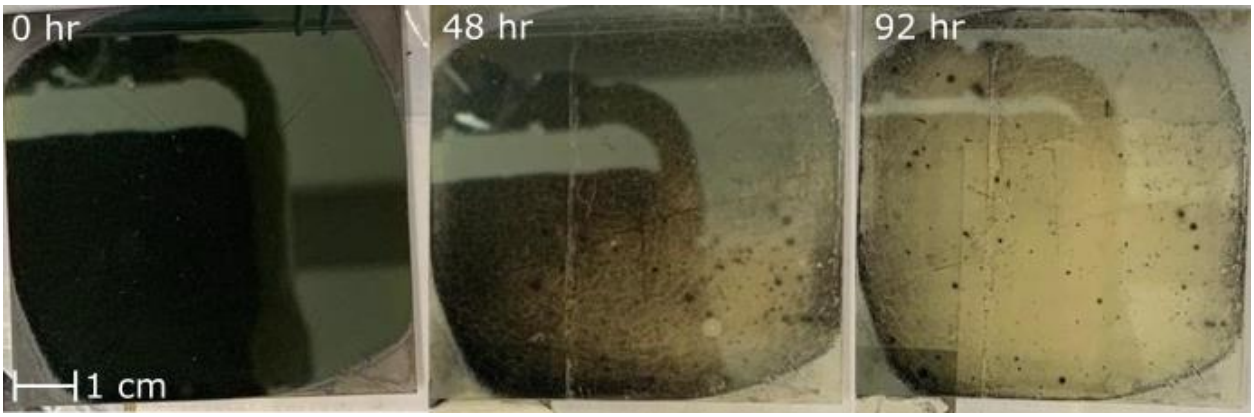

Figure 8.2: The same perovskite film with alumina barrier layer pictured after 0, 48, and 92 hours at 60% relative humidity and room temperature.

In this image, the gradient alumina barrier layer is thickest on the left, and we can see that the degradation begins from the side of the film which is thinner, with a notable pinhole in the lower-right of the 48hr image. The perovskite degraded much more quickly in this area, as the atmospheric water was able to more easily penetrate the thinner encapsulation. The perovskite films were kept in a consistent 60% relative humidity environment, and at room temperature until the entire area was degraded to lead iodide. Throughout this time the films were imaged at several-hour intervals, and these photos were subsequently separated into 50x50 pixel areas for use within our network. In comparison to the perovskite layer, calcium is a lot more reflective and the appearance of the film can change based on the ambient light conditions as can be seen in Figure 8.3. It also degrades to be completely clear, meaning that the background against which the image is taken will change the ease with which we can identify degraded samples. As the

area of the film that has degraded at a given point in time is critical to the WVTR calculation, we reduced the size of the square sub-samples to 20x20 pixels. WVTR is calculated through the formula:

$$\text{WVTR (g m}^{-2}\text{ d}^{-1}) = \text{n.}\delta_{Ca}\,\frac{M(H_2O)}{M(Ca)}\,\text{.h.}\,\frac{dA}{dt} \tag{23}$$

Where the property of interest for us is $\frac{dA}{dt}$, or the change in degraded area with respect to time. Conventionally this is taken as a constant, as it is challenging to accurately measure with any real precision. If we were able to track the degraded area in real-time, we could ascertain this rate of change to a significantly greater degree of precision, hence the smaller size of the squares in this phase of the project. In Equation 23, n is the molar equivalency between calcium and water (n=2), δCa is the density of calcium, in our evaporator this is a standard 1.55 g·cm^{-3}, M(H_2O)=18 g.mol^{-1} and M(Ca)=40 g.mol^{-1} are the molar densities of the two reacting species, and *h* is the thickness of the calcium film, measured to be 200nm through the evaporator's inbuilt quartz crystal microbalance.

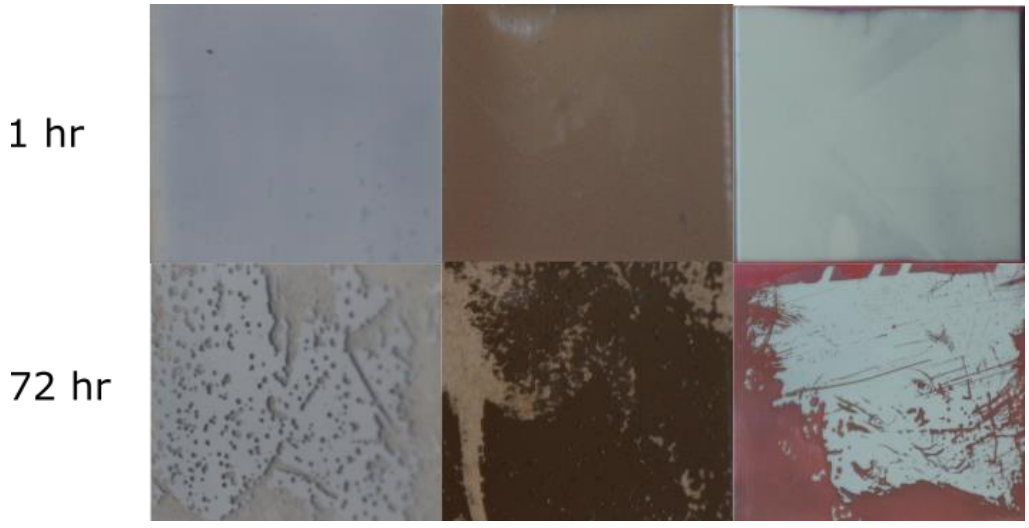

Figure 8.3: Three calcium films imaged at 1 and 72 hours. Varying ambient lighting conditions resulted in a much less reliable colour difference between intact and degraded films, so a red background layer was added for ease of identification of degraded and intact areas.

The variation in the appearance of the calcium film made it a lot harder to identify which of the squares contained degraded or intact film with a simple colour rule. However, because the calcium film degrades to be completely clear, we can add additional colour data to the background – making it easier to sort the training data into degraded and intact sub-samples, but also improving the ability of the CNN to distinguish the samples too. Red was selected as the background because Matlab breaks colour down into three channels, red, green, and blue. Intact perovskite films contain a lot of green, but a blue background would combine with degraded perovskite to produce green. Red is not as notable in the intact perovskite, and as yellow is a combination of red and green as far as the computer is concerned, the combination of degraded perovskite and red background (orange) still represents a relative lowering of the value in the green channel. This was intended to be useful for a potential future project wherein the two phases of this project were combined into a single algorithm able to identify visually degraded samples of both films. Through the nature of the CNN architecture within Matlab, we need to feed images of a specific size measured in pixels. To overcome this limitation, we mounted a fixed-focus camera to the top of the humidity chamber. This ensures that the ration of pixels:mm remains constant throughout all our tests.

8.2.3 Machine Learning Methods

Figure 8.4 is a generalized schematic of how the CNN integrates into our workflow. First the image of the film is broken down into small squares, in this case 50x50 pixels as we are looking at a perovskite film, whose location is indexed within Matlab. These squares are then passed to the CNN, the depth of which was determined through experimentation and based on the percentage of squares it classified correctly. The final layer of the CNN has a softmax activation function, which produces a set of relative probabilities that the square belongs to each class. The class with the highest probability is identified and reported, and a map of the film is created with degraded and intact areas labelled. In the case of the perovskite films it was necessary to include a class labelled "no film", as the spin-coating method used to deposit that material covers a circular area, leaving the corners of the glass slide with no film to degrade.

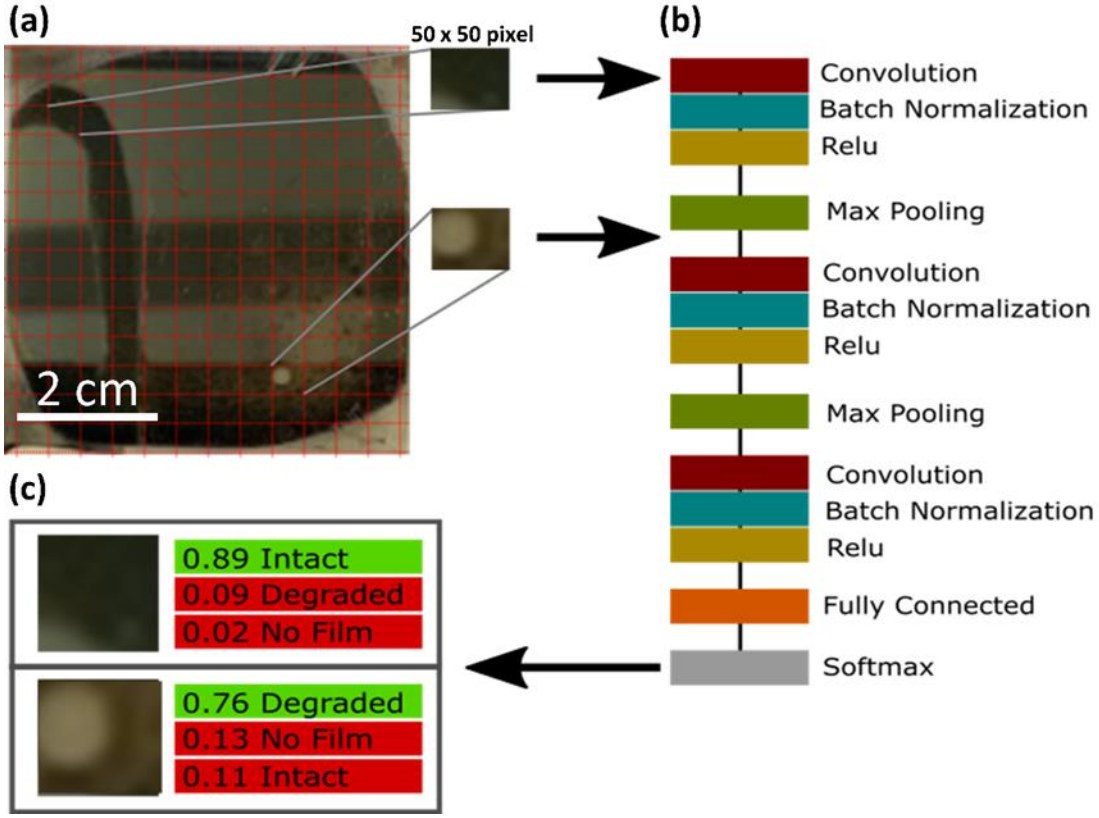

Figure 8.4: Schematic of the convolutional neural network developed for phase one of this work.

As could be seen in Figure 8.2 these corners remain visually distinct from even degraded film, but look more like degraded film than intact, and so the extra class was created to ensure that we were not overestimating the degraded area. Beyond providing mapping capabilities to our program, breaking the image of the film down into squares provides additional training data for the algorithm – hundreds of data points per film rather than just one. By shifting the window used to select which 50x50 pixel area is made into a square we could further increase the amount of training data and help ensure that the CNN was learning the colour change and edge detection that we can see are indicative of degraded films and no hidden features that may be counterproductive to our goal. These smaller images can also be processed much faster than larger ones, as our intent is to create a real-time monitoring system for precise barrier performance measurements.

8.3 Results and Discussion

Usually CNN's have to be comparatively deep networks, as they are often used for very complex image-recognition tasks. It should come as no surprise that correctly identifying the face of a specific individual under varying lighting conditions and with different cameras requires more computing power than detecting a simple colour change in images of specified size and under controlled conditions. That said, a network that is not complex enough will fail to accomplish even that. Using the same set of 10,000 50x50 pixel squares, four CNN's of increasing depth were trained and the performance of each network was measured using the number of squares that were sorted correctly into one of the three categories, degraded, intact, and no film. One-third of the squares in each training epoch were held back for validation, and a further 2000 squares were not used in training at all to provide a final testing result, as summarized below in TABLE IX.

TABLE IX: TRAINING, VALIDATION, AND TESTING PERFORMANCE OF FOUR CNN ALGORITHMS OF INCREASING DEPTH

Depth	3	5	10	25
Training	90.3%	93.8%	98.1%	98.5%
Validation	87.6%	88.1%	35.9%	11.2%
Testing	86.5%	84.9%	33.7%	12.4%

Here we defined depth as the number of convolution layers before the output layer, and we can see that the deeper networks actually performed worse during the testing and validation phases. This suggests that the networks were overfitting, and simply memorizing which squares belonged in which class – in a sense they were too powerful for the task at hand. In the end, we selected the three-layer network, as it performed slightly better when exposed to new samples compared to the five-layer network. Both classified the input images almost instantaneously, so both would have been appropriate for our application. Figure 8.5 shows how this algorithm worked in practice. On the right we have one of our gradient films imaged at several hour increments. Also indicated are the direction of increasing barrier thickness, and the area we

cropped to show the functionality of the CNN. The CNN was able to accurately track the progress of this degradation over 100 hours, starting with the defect in the lower-right of the film

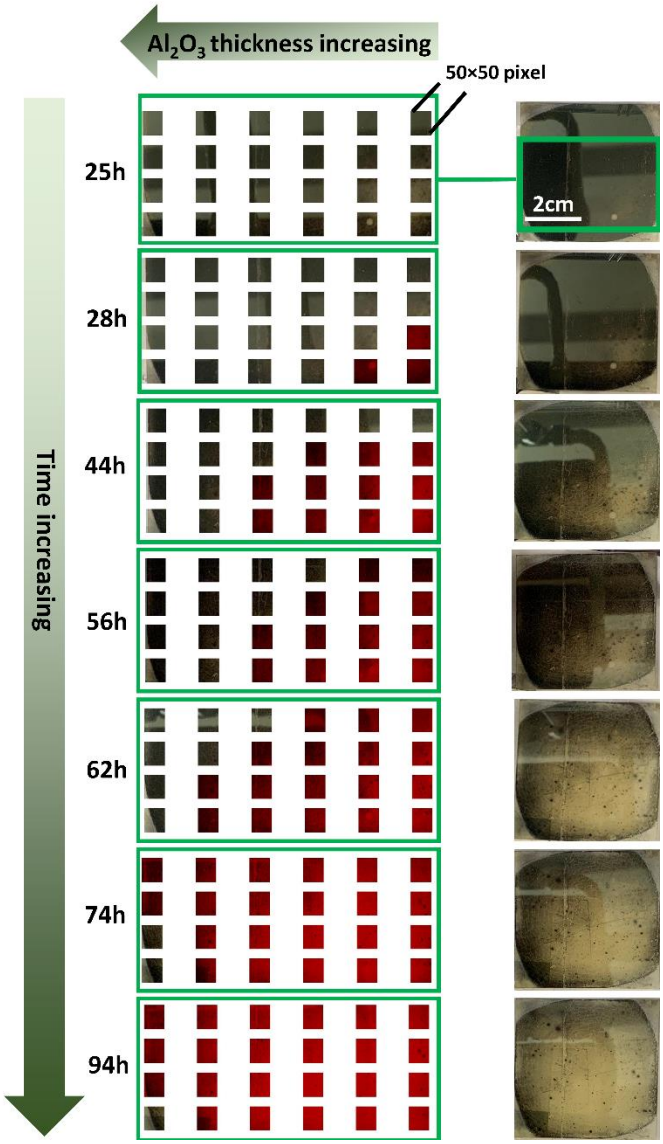

Figure 8.5: Final results of the perovskite degradation CNN, it can be seen that as time progresses the degraded squares marked in red show the same progression as the unedited film images.

at 28 hours. This defect induced increased degradation in the surrounding areas of the film, meaning we see a slight degradation gradient from bottom to top, perpendicular to the alumina thickness gradient. One important result from this experiment is that the alumina layer must be free of stress-induced cracking, a common problem in thick AP-SALD films. This occurs

because the varying layers apply stress to one another, much as with a composite at the macro-scale, but these stresses are small enough to not affect films below several hundred nanometres.

For the calcium samples, we used the same network architecture for two reasons. Firstly, the problem is not significantly different in terms of what we are asking the CNN to do, and secondly, we initially intended to create a hybrid algorithm that would process both kinds of films. As previously mentioned, the major change in the network itself was reducing the input image squares from 50x50 pixels, to 20x20 to more precisely track the degraded area, indicated

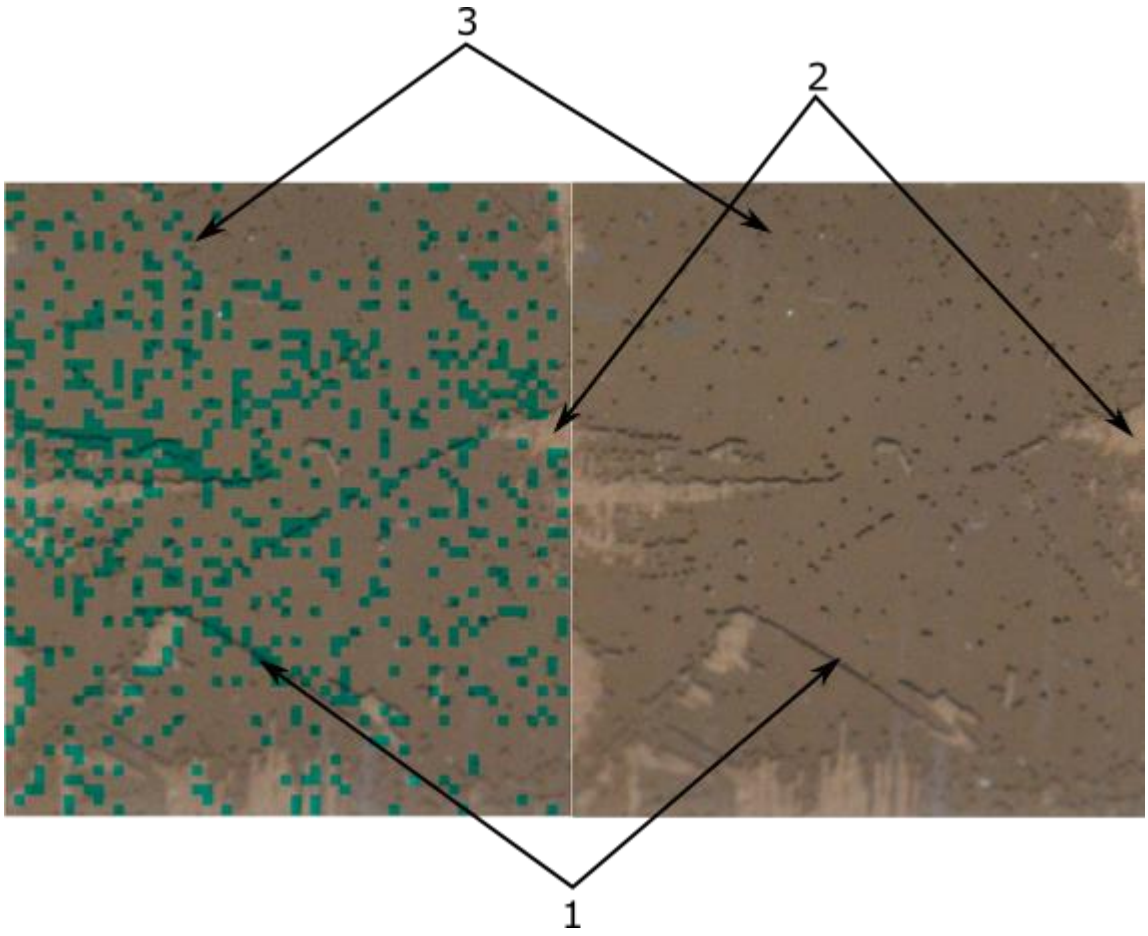

Figure 8.6: Processed vs original image of a calcium film trained with no added background. Green denotes areas classified as degraded. Notable features indicated, 1) is an angled line that was well captured by the algorithm, 2) is an open area of degraded film not captured by the algorithm, and 3) is an area of pinhole-defects that were also not well-resolved.

on Figure 8.6 in green. For the calcium WVTR network, initial performance was poor because

the difference between degraded and intact film was hard to detect even manually. The calcium film is so reflective that it takes on the colour of its environment much more readily than the perovskite films. Figure 8.6 has a classification accuracy of less than 20%; the network got very good at detecting the dark edges at the upper edge of the degraded areas (1), but would not detect any degradation within the exposed areas (2). We can also see that it did not detect many of the small pinholes throughout the film, particularly in the upper centre of the film (3), although some of these pinholes are so small it may not be justifiable to sort those squares into the "degraded" class regardless. Looking at the size of the green coloured pixels, it can be seen why it was so challenging to manually sort the 20x20 pixel areas into the correct classes. However, as the area of the degraded film is needed as accurately as possible for WVTR measurements, the small square size is necessary.

At this juncture, we added a red coloured background to our testing rig to improve both our ability to sort the squares manually and to provide greater contrast between the intact and degraded film. We re-trained the calcium WVTR CNN on over 120,000 20x20 pixel squares including the new background, split 6:3:3 between training, validation, and testing sets. As we can see in Figure 8.7, adding in the red background and re-training the model with the new data produced a much more accurate model, predicting the correct class 97.2% of the time compared to manual sorting. We can see from the image above, that the areas that are indicated as intact are very obviously calcium with no red background detected, save for small pinholes below the detection threshold. A possible solution for this would be to add in a third "partially-degraded" class to encompass these edge-cases, or to further reduce the size of the pixel squares being analysed. The network tends to assume that any square with more than 50% visible background is fully degraded, which was the metric I used when sorting the squares manually, and therefore

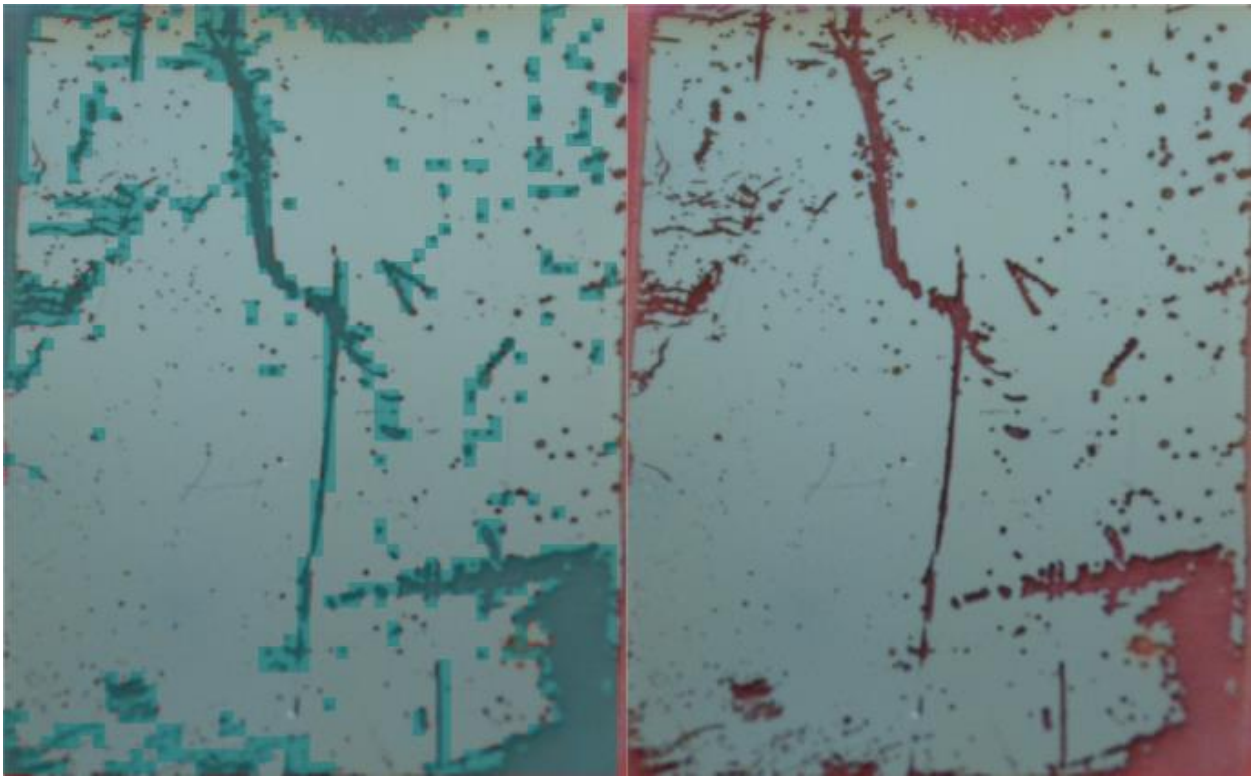

Figure 8.7: Re-trained calcium WVTR network using a red background to enhance visual contrast between degraded and intact films.

the future addition of a "partially-degraded" class should improve performance. The main advantage of this method over conventional image analysis programs such as ImageJ is that it is fully autonomous, but further work shall be done to improve the accuracy of our method and better account for the pinholes in films such as the one above. If we can approach the accuracy of conventional techniques with this autonomous method, then we will finally have a tool for ascertaining the validity of one of the core assumptions of the calcium WVTR test – that $\frac{dA}{dt}$ is linear. As it stands, because the pinholes are only smaller than our detection limit in the early stages of degradation, the accuracy of the autonomous area detection changes with time. In an attempt to broaden the applicability of the calcium WVTR network we tested combining squares with and without the added background, and while the performance when classifying films without the background was improved to 67%, the accuracy of the model when analysing films

with the background was reduced to 82.3%. This is likely because of the same difficulty with manually separating the dataset into the correct classes, a classic example of "garbage in, garbage out".

8.4 Conclusions

I developed two separate algorithms using computer vision techniques to assist us in combinatorial testing of encapsulation layers. The first iteration focused on autonomously mapping the degradation of perovskite films, and the second on calcium films used for WVTR testing. While there is significant work remaining to improve the precision of these models, they are both able to accurately determine if a sample is degraded or intact. This allows us to accelerate the design of new device stacks, by depositing films with graded properties and measuring which areas remain intact the longest – rather than depositing one film for each set of properties.

9 Conclusions and Next Stages

I have demonstrated in this work several means by which machine learning can be used to enhance materials science research in combination with our thin-film deposition technique. First, I lay the groundwork for my research by expanding the library of materials we can deposit within our lab, selecting a new organometallic precursor and allowing us to deposit copper and its oxides. Copper is the first metal that we have been able to produce in our lab, and its oxides have lower band gaps than our existing library of materials – and are already being used by other researchers to optimize their own device stacks.

Next, I developed and refined a method for measuring the resistance of our films in-situ, discovering along the way that by simply shifting the deposition temperature we can shift even simple zinc oxide films from semiconducting to metallic behaviour, likely through creation of oxygen defects in the film.

I then helped develop another in-situ measurement technique using a reflectance spectrometer to measure the thickness and band gap of our films. The Levenberg-Marquardt algorithm we used to fit the measured spectrum to an optical model suffered from poor initialization until I developed a machine learning method to extract an initial thickness estimate from the raw reflectance data. Over the course of several iterations, these machine-learning enhanced reflectance measurements predicted the thickness so accurately and with enough speed that we were able to implement a PI controller to control the rate of deposition in real-time.

Using a similar machine-learning method combined with spectroscopic ellipsometry I created a mechanism for measuring the properties of arbitrarily complex films accurately and quickly. Manually fitting films with multiple dopants to an optical model is a challenging and

time-consuming process. Conflicting information exists about the properties of undoped films, and so it was quickly apparent that we needed a method to approximate these properties first, retaining an expert in the loop to interpret how these approximations fit with our knowledge of the film. Great care had to be taken when developing the gaussian process regression model we used, applying transfer learning methods and ensuring that each iteration of the model was trained on the most accurate and comprehensive data possible. We verified film properties like the roughness and band gap of the film through the use of atomic force microscopy and absorbance spectroscopy and found that using optical-model estimates provided by the GPR model, we were able to measure these films using ellipsometry more accurately and within a fraction of the time compared to manual fitting.

Lastly, I applied computer vision techniques to assist with measuring the rate of degradation of water-sensitive films, and thereby the effectiveness of the barrier layers we deposited atop them using our AP-SALD system. Both perovskite and calcium films were imaged throughout the process of degrading. The pictures were sliced up into even squares, sorted into classes and used to train a convolutional neural network. The resulting convolutional neural networks were capable of accurately tracking which areas of the films were degraded and which were intact, allowing us to automate a testing process that can take several days to run to completion. Because the location of the degraded film is also tracked, we can now also do combinatorial testing where properties of the barrier films vary by location, providing more data in less time and resulting in significantly longer-lasting devices.

Through the course of this work, plans changed and new research directions were discovered. Initially, the plan had been to fully implement an ML-based system for controlling the AP-SALD - allowing us to simply specify a set of film properties and have the process

parameters selected for us. With the addition of the copper precursor the range of possible materials that we can produce increased massively. As mentioned previously it is capable of producing metallic copper, as well as cupric and cuprous oxide individually and as a mixture of any number of phases. Using these materials as components in a multi-oxide film, like Cu:ZnO or Cu:AZO increased the range of possible properties massively compared to the original intended data-space involving just Al and Zn oxides. Accordingly, this increased the complexity of accurately finding the film properties using our laboratory equipment by a similar magnitude, and the amount of time and data needed for the previously-described sections of the report. Unfortunately, there was not sufficient time to accomplish my original goals, but in Section 9.1 I will expand upon my plans for how this might be accomplished. In Section 9.2 I will explain an exciting new application for our AP-SALD films, especially highly-doped films like those we can create with these new copper-based materials – εNZ materials which display nonlinear optical properties. These properties allow for the production of novel devices like optical transistors, which can be made much smaller than the current limit of FET designs, being much simpler in structure and less susceptible to temperature as there is no current flowing through the device.

9.1 Machine-Learning Integrated Control System

The first problem we aim to solve with machine learning is the prediction of what the material properties of a film are going to be, once the deposition parameters are entered into the AP-SALD/CVD system. As previously mentioned, this is a complex multivariable problem, each deposition parameter can have multiple effects on the final film, and the interrelations between even just two parameters and the final film properties would be incredibly complex to try and characterize analytically. Since we have eight deposition parameters and four independent

variables, it may be necessary to alter the dataset so that the ML model is a simple as possible. The use of autoencoders and principal component analysis will show which parameters or combinations of parameters actually have an effect on the material properties we are looking at, hopefully allowing us to further reduce the number of input dimensions, beyond what is possible with manual experimentation. We have already made significant strides in this direction with the three measurement techniques described in the preceding report, and the compilation of the film property database used for training the ellipsometry ML model.

Initially we had conceived of using a recurrent neural network for this task, as we would be able to adjust and update the film deposition parameters in real-time. However, essentially having to repeat the development of the ML model presented in Chapter 7 for each time-step while varying the deposition parameters throughout to encompass the effects of each potential control variable is a significant proposition that would monopolize the lab-scale SALD for years. Much simpler would be to assume that the process can be controlled using only the parameters used in the beginning, and that they remain constant throughout the deposition. This assumption will allow us to use the database we compiled for the ellipsometry ML model as a starting point, but instead of including the ellipsometry data as an input and the material properties as an output we would be using the properties as an input and the deposition parameters as an output. We already have all of this data, however there may be multiple sets of outputs that are equally viable based on the desired inputs. One solution to this would be to place constraints on what the model is allowed to produce, such as limiting what materials and temperature ranges can be used. This would be easier to program, and also give the user greater control over the deposition process. However, this could still be accomplished in conjunction with the second potential solution, the use of a generative adversarial network (GAN).

A GAN is a unique ML formed of two networks working in opposition, hence the ‘adversarial’ part of their name [257], [258]. The first network is called a ‘generator’, and its job is to learn how to create convincing but artificial outputs, given a set of inputs. The second is referred to as a ‘discriminator’, and its job is to determine if its inputs are real or generated by the first network. GAN’s can be comprised of any kind of ANN but are most commonly a paired CNN/inverse CNN. GAN/CNN pairs are often used to generate artificial images which are convincing to the human eye, such as *Edmond de Belamy,* the first work of art created by an AI which sold for $432,500 in 2018. While later image generation techniques have surpassed the performance of GAN’s when creating artwork, they are still well-designed to solve any problem which can be solved both forwards and in reverse. This is particularly useful when there is an imbalance between the number of potential inputs and outputs, or as in our case when we want to generate one workable solution where many potential ones exist.

GAN’s are comparatively complex to train, as you must create two separate but interdependent networks, each of which may be many layers deep, and this has slowed their adoption outside computer science. A common failing of GAN’s is called modal collapse, where the generator only produces results around a single point which has fooled the discriminator in the past, where the data may have a much broader range. The first step to training a GAN is to train the discriminator as you would any ANN, so it can recognize the inputs. Next, randomly modify an input sample, and train the generator to reconstruct the original input, learning latent variables from the input which may not be apparent upon initial observation. Feed both the reconstructed and original outputs into the discriminator, which will then determine which input is real. The classification error of the discriminator is then backpropagated into both networks, the weights in the discriminator are adjusted to minimize this value while the generator is

adjusted to maximize it. While GAN's may be most commonly used in image processing tasks, it does not mean they cannot be applied effectively to materials science. Relating the microstructure of a material to its properties is a comparatively simple problem to solve, but trying to predict a materials microstructure from its properties is a much more complex task [259]. A GAN forms the backbone of the computational process outlined, since the material properties are learned from observing the microstructures without pre-processing, the most information possible is preserved before being fed into the generator. The latent variables learned by the generator are used as design targets, and the algorithm generates a microstructure with those properties. This method is both faster and more informative than current microstructure characterization and reconstruction techniques. In a similar manner, because multiple sets of deposition parameters could provide the same material properties, the ability of a GAN to solve this kind of inverse problem is appealing.

9.2 Discovery of New ε-Near Zero Materials

Through the development of our ellipsometry model in Chapter 7 we noted that degeneratively-doped semiconductors such as AZO can exhibit metal-like properties represented by the Drude component of the multi-oscillator model [260]. AZO in particular is notable for how easy it is to tailor the dielectric permittivity of the material in the infrared and near-infrared [261], an effect characterized by the increase in carriers and switch to metallic behaviour like we saw in our ZnO films in Chapter 0. These carriers can come from many sources, but as we suggested in that chapter oxygen defects within the film lattice are a significant contributor, and one that our AP-SALD system is ideally suited to manipulate. In addition, while conventional ALD has been used to deposit AZO films with very high defect concentrations [262], as we know, this is a much slower process than our AP-SALD/CVD machine. ε-near zero materials

(ENZ) are still very experimental, but they have several applications in advanced computing and communications where the rapid fabrication capacity of an AP-SALD system will be integral to rapid commercial development.

One particularly useful application for these ε-near zero materials is in the development of optical transistors. It is becoming increasingly difficult to manufacture smaller metal-oxide semiconductor field effect transistors (MOSFET) because they are becoming so small that electrons may tunnel through the device instead of allowing it to switch effectively, issues with effectively cooling such small devices, and that's not to mention the difficulty of reliably manufacturing devices on the near-atomic scale. ENZ materials have a unique property where their optical properties change when illuminated around the wavelength where their dielectric permittivity (ε) is near or equal to zero, which occurs at the plasma frequency of the material as modelled by the Drude optical model. In practice a first, "switch" pulse of light at the plasma frequency interacts with free electrons within the material, polarizing the surface plasmons and scattering light for a given length of time, the "probe" pulse. The switch-probe pulse interactions allow us to construct a device similar in function to a transistor but using optical pulses to transmit data and not electrons, illustrated below in Figure 9.1. In a conventional MOSFET we

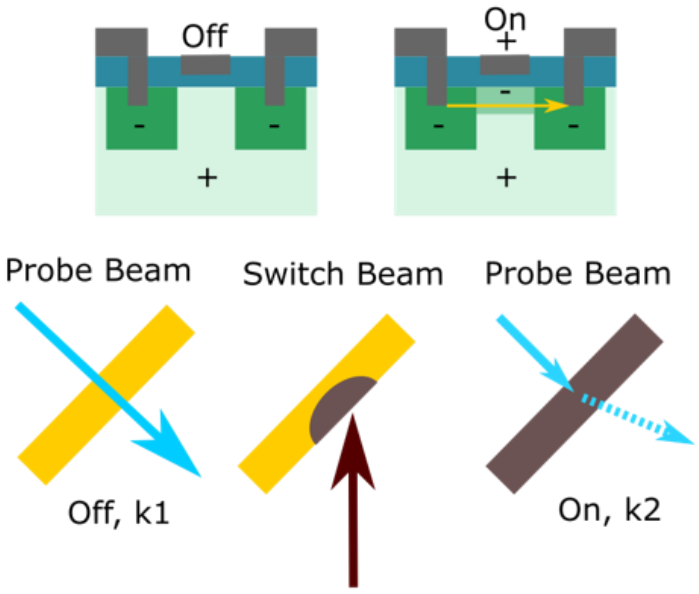

Figure 9.1: Metal-oxide semiconductor field effect transistor (top) vs. ENZ optical transistor (bottom)

create a channel through which electrons can pass by applying a voltage to the switch terminal, and the field effect creates a channel of negatively-charged semiconductor between the two terminals. In an optical transistor the probe beam passes through the device in one way when it's off, then the switch beam polarizes the material, changing the value of the extinction coefficient, bending the second probe pulse in a different manner. By changing where the detector for the probe pulse is, we can ensure we only detect one of these beams resulting in a signal that is either on or off based on the length of time since the last switch pulse. AZO is a particularly good ENZ material for several reasons, firstly its band gap is non-parabolic, so its plasma frequency is more susceptible to electron temperature and therefore, the change in optical properties is greater for a given pump pulse. In addition, AZO has a low free electron density so the electron heat capacity is lower, further amplifying this effect in comparison to metals which are often also considered for ENZ applications. Lastly, the plasma frequency of AZO is often in the near infra-red, where

many commercial pulsed lasers are available, meaning that we would not need to buy specialist lasers for the switch or probe pulses.

There are challenges in both the realms of material science and commercial scale-up that must be solved before optical transistors can replace MOSFETs in computers. Through utilizing our experience of creating optical models for ellipsometry, and our ability to tailor film properties and doping levels through altering our AP-SALD/CVD deposition parameters we are well-placed for optimizing these materials and then producing them at scale. Given our library of precursors, capacity for creating gradient films, and existing material property database we can use ML methods to tune the plasma frequency and carrier density of these films, and test them in collaboration with other groups already working in the field. Accelerating the adoption of optical transistors will allow for the current rapid pace of increasing computer power to continue for much longer than with conventional MOSFET's, and presents an intriguing opportunity for an additional commercial spin-off for our lab.

Letters of Copyright Permission

Several figures used in the production of this thesis have been previously published, some under copyright protection and others under a Creative Commons 4.0 (CC by 4.0) license, the terms of which can be found at https://creativecommons.org/licenses/by/4.0/legalcode. The disclaimer of warranties for this license is as follows:

Section 5 – Disclaimer of Warranties and Limitation of Liability.

a. *Unless otherwise separately undertaken by the Licensor, to the extent possible, the Licensor offers the Licensed Material as-is and as-available, and makes no representations or warranties of any kind concerning the Licensed Material, whether express, implied, statutory, or other. This includes, without limitation, warranties of title, merchantability, fitness for a particular purpose, non-infringement, absence of latent or other defects, accuracy, or the presence or absence of errors, whether or not known or discoverable. Where disclaimers of warranties are not allowed in full or in part, this disclaimer may not apply to You.*

b. *To the extent possible, in no event will the Licensor be liable to You on any legal theory (including, without limitation, negligence) or otherwise for any direct, special, indirect, incidental, consequential, punitive, exemplary, or other losses, costs, expenses, or damages arising out of this Public License or use of the Licensed Material, even if the Licensor has been advised of the possibility of such losses, costs, expenses, or damages. Where a limitation of liability is not allowed in full or in part, this limitation may not apply to You.*

c. *The disclaimer of warranties and limitation of liability provided above shall be interpreted in a manner that, to the extent possible, most closely approximates an absolute disclaimer and waiver of all liability.*

In accordance with this license, the caption of each relevant image includes the primary author of the paper from which the image originates, as well as the address at which the original article may be found. As pseudonyms are acceptable under this license, I shall refer to myself by my chosen name, even as this may differ from the name listed at time of publication of the original material.

In addition, some material was taken from the paper I co-authored "*Nanoscale Film Thickness Gradients Printed in Open Air by Spatially Varying Chemical Vapor Deposition*", Specifically Figure 8.4 and Figure 8.5, the license for which is copied below.

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS
May 03, 2023

This Agreement between University of Waterloo -- Olivia Marchione ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	5541420301794
License date	May 03, 2023
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Advanced Functional Materials
Licensed Content Title	Nanoscale Film Thickness Gradients Printed in Open Air by Spatially Varying Chemical Vapor Deposition
Licensed Content Author	Kevin P. Musselman, David Muñoz-Rojas, Mustafa Yavuz, et al
Licensed Content Date	Jun 10, 2021
Licensed Content Volume	31
Licensed Content Issue	31
Licensed Content Pages	14
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Electronic
Portion	Figure/table
Number of figures/tables	2
Will you be translating?	No
Title	¬Leveraging Atmospheric-Pressure Spatial Atomic Layer Deposition and Machine Learning for Nanomaterial and Device Design
Institution name	University of Waterloo
Expected presentation date	Jun 2023

Portions	Figure 7e) from the main text, as well as figure S9 from the supporting information, as these represent my contribution to the work.
Requestor Location	University of Waterloo 196 Westmount Rd N Unit 505 Waterloo, ON N2L 3G66 Canada Attn: University of Waterloo
Publisher Tax ID	EU826007151
Total	0.00 CAD

Terms and Conditions

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a"Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your RightsLink account (these are available at any time at http://myaccount.copyright.com).

Terms and Conditions

- The materials you have requested permission to reproduce or reuse (the "Wiley Materials") are protected by copyright.
- You are hereby granted a personal, non-exclusive, non-sub licensable (on a stand-alone basis), non-transferable, worldwide, limited license to reproduce the Wiley Materials for the purpose specified in the licensing process. This license, **and any CONTENT (PDF or image file) purchased as part of your order,** is for a one-time use only and limited to any maximum distribution number specified in the license. The first instance of republication or reuse granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before the end date may be distributed thereafter). The Wiley Materials shall not be used in any other manner or for any other purpose, beyond what is granted in the license. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley

publication in your use of the Wiley Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Wiley Material. Any third party content is expressly excluded from this permission.

- With respect to the Wiley Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Wiley Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Wiley Materials without the prior permission of the respective copyright owner.**For STM Signatory Publishers clearing permission under the terms of the STM Permissions Guidelines only, the terms of the license are extended to include subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and does not involve the separate exploitation of the permitted figures or extracts,** You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Wiley Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.

- The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto

- NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

- WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

- You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

- IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

- Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

- The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

- This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

- Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.

- These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This

Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

- In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

- WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

- This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

- This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

WILEY OPEN ACCESS TERMS AND CONDITIONS

Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

The Creative Commons Attribution License

The Creative Commons Attribution License (CC-BY) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-

Creative Commons Attribution Non-Commercial License

The Creative Commons Attribution Non-Commercial (CC-BY-NC)License permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.(see below)

Creative Commons Attribution-Non-Commercial-NoDerivs License

References

H. Zhang and S. K. Moon, "Reviews on machine learning approaches for process optimization in noncontact direct ink writing," *ACS Appl Mater Interfaces*, vol. 13, no. 45, pp. 53323–53345, Nov. 2021, doi: 10.1021/ACSAMI.1C04544.

[2] H. Zhang, S. K. Moon, and T. H. Ngo, "Hybrid Machine Learning Method to Determine the Optimal Operating Process Window in Aerosol Jet 3D Printing," *ACS Appl Mater Interfaces*, p. acsami.9b02898, May 2019, doi: 10.1021/acsami.9b02898.

[3] S. Zeng, G. Li, Y. Zhao, R. Wang, and J. Ni, "Machine Learning Aided Design of Materials with Target Elastic Properties," *J. Phys. Chem. C*, vol. 123, no. 8, Feb. 2019, doi: 10.1021/acs.jpcc.9b01045.

[4] S. P. Ong, "Accelerating materials science with high-throughput computations and machine learning," *Comput Mater Sci*, vol. 161, pp. 143–150, Apr. 2019, doi: 10.1016/J.COMMATSCI.2019.01.013.

[5] J. M. Rickman, T. Lookman, and S. V. Kalinin, "Materials Informatics: From the Atomic-Level to the Continuum," *Acta Mater*, Jan. 2019, doi: 10.1016/J.ACTAMAT.2019.01.051.

[6] O. Olesegum, R. Noraas, N. Somanath, and M. Giering, "Structural Material Property Tailoring Using Deep Neural Networks," *ArXiv*, pp. 1–12, doi: 10.2514/6.2019-1703.

[7] Y. Liu, T. Zhao, W. Ju, and S. Shi, "Materials discovery and design using machine learning," *Journal of Materiomics*, vol. 3, no. 3, pp. 159–177, Sep. 2017, doi: 10.1016/J.JMAT.2017.08.002.

[8] K. Choudhary, M. Bercx, J. Jiang, R. Pachter, and D. Lamoen, "Accelerated Discovery of Efficient Solar-cell Materials using Quantum and Machine-learning Methods," *Chem. Mater*, vol. 31, no. 15, pp. 5900–5908, Jul. 2019.

[9] G. Dogan *et al.*, "Bayesian machine learning for efficient minimization of defects in ALD passivation layers," *ACS Appl Mater Interfaces*, vol. 13, no. 45, pp. 54503–54515, Nov. 2021, doi: 10.1021/acsami.1c14586.

[10] Y. Ding, Y. Zhang, Y. M. Ren, G. Orkoulas, and P. D. Christofides, "Machine learning-based modeling and operation for ALD of SiO2 thin-films using data from a multiscale CFD simulation," *Chemical Engineering Research and Design*, vol. 151, pp. 131–145, Nov. 2019, doi: 10.1016/J.CHERD.2019.09.005.

[11] Y. Ding, Y. Zhang, K. Kim, A. Tran, Z. Wu, and P. D. Christofides, "Microscopic Modeling and Optimal Operation of Thermal Atomic Layer Deposition," *Chemical Engineering Research and Design*, Mar. 2019, doi: 10.1016/J.CHERD.2019.03.004.

[12] H. O. Pierson, "Fundamentals of Chemical Vapor Deposition," in *Handbook of Chemical Vapor Deposition (CVD)*, Norwich, New York: Noyes Publications, 1999, pp. 36–67. doi: 10.1016/B978-081551432-9.50005-X.

[13] T. Suntola and J. Antson, "Method for producing compound thin films," US4058430A, Nov. 25, 1975 Accessed: Jun. 22, 2019. [Online]. Available: https://patents.google.com/patent/US4058430A/en

[14] T. Suntola, A. Pakkala, and S. Lindfors, "Apparatus for performing growth of compound thin films," US4389973A, Dec. 11, 1983 Accessed: Feb. 25, 2019. [Online]. Available: https://patents.google.com/patent/US4389973A/en

[15] T.-Y. Chiang, C.-L. Dai, and D.-M. Lian, "Influence of growth temperature on the optical and structural properties of ultrathin ZnO films," *J Alloys Compd*, vol. 509, no. 18, pp. 5623–5626, May 2011, doi: 10.1016/J.JALLCOM.2011.02.093.

[16] K.-Y. Yen *et al.*, "On the Characteristics of Zinc Oxide Films Grown on (11-20) Sapphire Substrates by Atomic Layer Deposition Using Diethylzinc and Nitrous Oxide," in *ECS Transactions*, ECS, May 2009, pp. 843–853. doi: 10.1149/1.3122138.

[17] T. Tynell and M. Karppinen, "Atomic layer deposition of ZnO: A review," *Semicond Sci Technol*, vol. 29, no. 4, Jan. 2014, doi: 10.1088/0268-1242/29/4/043001.

[18] A. Perrotta, J. Pilz, A. Milella, and A. M. Coclite, "Opto-chemical control through thermal treatment of plasma enhanced atomic layer deposited ZnO: An in situ study," *Appl Surf Sci*, Mar. 2019, doi: 10.1016/J.APSUSC.2019.03.122.

[19] S. Roy and S. Basu, "Improved zinc oxide film for gas sensor applications," *Bulletin of Materials Science*, vol. 25, no. 6, pp. 513–515, Nov. 2002, doi: 10.1007/BF02710540.

[20] Z. Yu *et al.*, "Development of photocatalytically active heterostructured MnO/ZnO and CuO/ZnO films via a solution precursor plasma spray process," *Surf Coat Technol*, Feb. 2019, doi: 10.1016/J.SURFCOAT.2019.02.053.

[21] B. Hussain *et al.*, "Electron Affinity and Bandgap Optimization of Zinc Oxide for Improved Performance of ZnO/Si Heterojunction Solar Cell Using PC1D Simulations," *Electronics (Basel)*, vol. 8, no. 2, p. 238, Feb. 2019, doi: 10.3390/electronics8020238.

[22] Y. Chen, Y. Hu, Q. Meng, H. Yan, W. Shuai, and Z. Zhang, "Natively textured surface of Ga-doped ZnO films electron transporting layer for perovskite solar cells: further performance analysis from device simulation," *Journal of Materials Science: Materials in Electronics*, pp. 1–11, Jan. 2019, doi: 10.1007/s10854-019-00766-7.

[23] M. F. Mohamad Noh *et al.*, "The architecture of the electron transport layer for a perovskite solar cell," *J Mater Chem C Mater*, vol. 6, no. 4, pp. 682–712, Nov. 2018, doi: 10.1039/c7tc04649a.

[24] A. S. Asundi, J. A. Raiford, and S. F. Bent, "Opportunities for Atomic Layer Deposition in Emerging Energy Technologies," *ACS Energy Lett*, p. acsenergylett.9b00249, Mar. 2019, doi: 10.1021/acsenergylett.9b00249.

[25] S. Lu, H. Wang, J. Zhou, X. Wu, and W. Qin, "Atomic layer deposition of ZnO on carbon black as nanostructured anode materials for high-performance lithium-ion batteries," *Nanoscale*, vol. 9, no. 3, pp. 1184–1192, Jan. 2017, doi: 10.1039/C6NR07868K.

[26] B. S. Witkowski, "Applications of ZnO Nanorods and Nanowires-A Review," *Acta Physica Polonica Series A*, vol. 134, no. 6, pp. 1226–1246, Dec. 2018, doi: 10.12693/APhysPolA.134.1226.

[27] A. Rezk and I. Saadat, *High-Performance ALD Al-doped ZnO Thin Film Transistors Grown on Flexible Substrates*. 2019, pp. 1–1. doi: 10.1016/j.cub.2015.10.018.

[28] A. Abliz *et al.*, "Investigation on the electrical performances and stability of W-doped ZnO thin-film transistors," *Mater Sci Semicond Process*, vol. 95, pp. 54–58, Jun. 2019, doi: 10.1016/J.MSSP.2019.01.027.

[29] M. Wang *et al.*, "High-Performance Flexible ZnO Thin-Film Transistors by Atomic Layer Deposition," *IEEE Electron Device Letters*, vol. PP, no. c, pp. 1–1, Jan. 2019, doi: 10.1109/led.2019.2895864.

[30] K. Mistry, A. Jones, M. Kao, T. W.-K. Yeow, M. Yavuz, and K. P. Musselman, " In-situ observation of nucleation and property evolution in films grown with an atmospheric pressure spatial atomic layer deposition system ," *Nano Express*, vol. 1, no. 1, p. 010045, Jun. 2020, doi: 10.1088/2632-959x/ab976c.

[31] A. Jones, K. Mistry, M. Kao, A. Shahin, M. Yavuz, and K. P. Musselman, "In-situ spatial and temporal electrical characterization of ZnO thin films deposited by atmospheric pressure chemical vapour deposition on flexible polymer substrates," *Scientific Reports 2020 10:1*, vol. 10, no. 1, pp. 1–10, Nov. 2020, doi: 10.1038/s41598-020-76993-4.

[32] K. Mistry, "Applications of Reflectometry Towards the Development of MEMS Gas Sensors," May 2021, Accessed: Nov. 05, 2022. [Online]. Available: https://uwspace.uwaterloo.ca/handle/10012/17031

[33] V. H. Nguyen, U. Gottlieb, A. Valla, D. Muñoz, D. Bellet, and D. Muñoz-Rojas, "Electron tunnelling through grain boundaries in transparent conductive oxides and implications for electrical conductivity: the case of ZnO:Al thin films," *Mater Horiz*, vol. 5, no. 4, pp. 715–726, Jul. 2018, doi: 10.1039/C8MH00402A.

[34] A. Sekkat *et al.*, "Open-air printing of Cu2O thin films with high hole mobility for semitransparent solar harvesters," *Communications Materials 2021 2:1*, vol. 2, no. 1, pp. 1–10, Jul. 2021, doi: 10.1038/s43246-021-00181-8.

[35] A. H. Alshehri *et al.*, "Nanoscale Film Thickness Gradients Printed in Open Air by Spatially Varying Chemical Vapor Deposition," *Adv Funct Mater*, vol. 31, no. 31, p. 2103271, Aug. 2021, doi: 10.1002/ADFM.202103271.

[36] C. Armstrong, L. V. Delumeau, D. Muñoz-Rojas, A. Kursumovic, J. MacManus-Driscoll, and K. P. Musselman, "Tuning the band gap and carrier concentration of titania films grown by spatial atomic layer deposition: a precursor comparison," *Nanoscale Adv*, vol. 3, no. 20, pp. 5908–5918, Oct. 2021, doi: 10.1039/D1NA00563D.

[37] K. Mistry *et al.*, "Highly Sensitive Self-Actuated Zinc Oxide Resonant Microcantilever Humidity Sensor," *Nano Lett*, vol. 22, no. 8, pp. 3196–3203, Apr. 2022, doi: 10.1021/ACS.NANOLETT.1C04378/ASSET/IMAGES/LARGE/NL1C04378_0004.JPEG.

[38] L. V. Delumeau *et al.*, "Effectiveness of antiviral metal and metal oxide thin-film coatings against human coronavirus 229E," *APL Mater*, vol. 9, no. 11, p. 111114, Nov. 2021, doi: 10.1063/5.0056138.

[39] J. Resende *et al.*, "Planar and Transparent Memristive Devices Based on Titanium Oxide Coated Silver Nanowire Networks with Tunable Switching Voltage," *Small*, vol. 17, no. 21, p. 2007344, May 2021, doi: 10.1002/SMLL.202007344.

[40] G. V Naik *et al.*, "Oxides and nitrides as alternative plasmonic materials in the optical range," *IEDM Technical Digest. International*, vol. 331, no. 310, p. 18, Aug. 2011, doi: 10.48550/arxiv.1108.0993.

[41] A. K. Pradhan *et al.*, "Extreme tunability in aluminum doped Zinc Oxide plasmonic materials for near-infrared applications," *Scientific Reports 2014 4:1*, vol. 4, no. 1, pp. 1–6, Sep. 2014, doi: 10.1038/srep06415.

[42] "Hazardous Substances Data Bank (HSDB) : 929 - PubChem." https://pubchem.ncbi.nlm.nih.gov/source/hsdb/929

[43] E. Llc, "Ereztech LLC SN6779 Safety Data Sheet", [Online]. Available: https://ereztech.com

[44] C. Pan *et al.*, "Sprayable copper and copper–zinc nanowires inks for antiviral surface coating," *RSC Adv*, vol. 12, no. 10, pp. 6093–6098, 2022, doi: 10.1039/D1RA08755J.

[45] P. Merkl, S. Long, G. M. McInerney, and G. A. Sotiriou, "Antiviral activity of silver, copper oxide and zinc oxide nanoparticle coatings against sars-cov-2," *Nanomaterials*, vol. 11, no. 5, p. 1312, May 2021, doi: 10.3390/nano11051312.

[46] I. Rani *et al.*, "Potential molecular mechanisms of zinc- and copper-mediated antiviral activity on COVID-19," *Nutrition Research*, Jun. 2021, doi: 10.1016/j.nutres.2021.05.008.

[47] G. Pullangott, U. Kannan, G. S., D. V. Kiran, and S. M. Maliyekkal, "A comprehensive review on antimicrobial face masks: An emerging weapon in fighting pandemics," *RSC Advances*, vol. 11, no. 12. Royal Society of Chemistry, pp. 6544–6576, Jan. 20, 2021. doi: 10.1039/d0ra10009a.

[48] G. Borkow and J. Gabbay, "Putting copper into action: copper-impregnated products with potent biocidal activities," *The FASEB Journal*, vol. 18, no. 14, pp. 1728–1730, Nov. 2004, doi: 10.1096/fj.04-2029fje.

[49] R. L. Z. Hoye *et al.*, "Perspective: Maintaining surface-phase purity is key to efficient open air fabricated cuprous oxide solar cells," *APL Mater*, vol. 3, no. 2, p. 020901, Feb. 2015, doi: 10.1063/1.4913442.

[50] A. H. Hammad and M. Sh. Abdel-wahab, "Photocatalytic activity in nanostructured zinc oxide thin films doped with metallic copper," *Physica B Condens Matter*, vol. 646, p. 414352, Dec. 2022, doi: 10.1016/J.PHYSB.2022.414352.

[51] A. P. Cabello, M. A. Ulla, and J. M. Zamaro, "In situ growth of nanostructured copper and zinc mixed oxides on brass supports as efficient microreactors for the catalytic CO oxidation," *Journal of Materials Science 2022*, pp. 1–13, Jul. 2022, doi: 10.1007/S10853-022-07391-6.

[52] L. Pan *et al.*, "Boosting the performance of Cu2O photocathodes for unassisted solar water splitting devices," *Nat Catal*, vol. 1, no. 6, pp. 412–420, Jun. 2018, doi: 10.1038/s41929-018-0077-6.

[53] P. G. Gordon, A. Kurek, and S. T. Barry, "Trends in Copper Precursor Development for CVD and ALD Applications," *ECS Journal of Solid State Science and Technology*, vol. 4, no. 1, pp. N3188–N3197, Dec. 2015, doi: 10.1149/2.0261501JSS/XML.

[54] J. A. T. Norman, M. Perez, S. E. Schulz, and T. Waechtler, "New precursors for CVD copper metallization," *Microelectron Eng*, vol. 85, no. 10, pp. 2159–2163, Oct. 2008, doi: 10.1016/J.MEE.2008.05.036.

[55] S. W. Kang, J. Y. Yun, and Y. H. Chang, "Growth of Cu metal films at room temperature using catalyzed reactions," *Chemistry of Materials*, vol. 22, no. 5, pp. 1607–1609, Mar. 2010, doi: 10.1021/CM902294E/ASSET/IMAGES/LARGE/CM-2009-02294E_0003.JPEG.

[56] R. Solanki and B. Pathangey, "Atomic Layer Deposition of Copper Seed Layers," *Electrochemical and Solid-State Letters*, vol. 3, no. 10, pp. 479–480, Aug. 2000.

[57] Z. Zhong, X. Wang, J. Ding, and N. Yuan, "Nanometer-thick copper films grown by thermal atomic layer deposition," *Thin Solid Films*, vol. 589, pp. 673–680, Aug. 2015, doi: 10.1016/J.TSF.2015.06.053.

[58] J. Huo, R. Solanki, and J. Mcandrew, "Characteristics of copper films produced via atomic layer deposition," *J Mater Res*, vol. 17, no. 9, pp. 2394–2398, Sep. 2002, doi: 10.1557/JMR.2002.0350.

[59] A. Sekkat *et al.*, "Selective spatial atomic layer deposition of Cu, Cu2O, and CuO thin films in the open air: reality or fiction?," *Mater Today Chem*, vol. 29, p. 101431, Apr. 2023, doi: 10.1016/J.MTCHEM.2023.101431.

[60] G. Papadimitropoulos and D. Davazoglou, "Copper Films Deposited by Hot-Wire CVD and Direct Liquid Injection of CupraSelect," *Chemical Vapor Deposition*, vol. 13, no. 11, pp. 656–662, Nov. 2007, doi: 10.1002/CVDE.200706621.

[61] T. Waechtler *et al.*, "Copper Oxide Films Grown by Atomic Layer Deposition from Bis(tri-n-butylphosphane)copper(I)acetylacetonate on Ta, TaN, Ru, and SiO2," *J Electrochem Soc*, vol. 156, no. 6, p. H453, Apr. 2009, doi: 10.1149/1.3110842/XML.

[62] E. Machado *et al.*, "Interaction of copper organometallic precursors with barrier layers of Ti, Ta and W and their nitrides: A first-principles molecular dynamics study," *J Mol Model*, vol. 13, no. 6–7, pp. 861–864, Jul. 2007, doi: 10.1007/S00894-007-0187-6/FIGURES/3.

[63] F. M. Li *et al.*, "Low temperature (< 100 °c) deposited P-type cuprous oxide thin films: Importance of controlled oxygen and deposition energy," *Thin Solid Films*, vol. 520, no. 4, pp. 1278–1284, Dec. 2011, doi: 10.1016/J.TSF.2011.04.192.

[64] "Copper - Electrical Resistivity and Electrical Conductivity." https://www.periodic-table.org/Copper-electrical-resistivity/

[65] S. Privitera, E. Rimini, and R. Zonca, "Amorphous-to-crystal transition of nitrogen- and oxygen-doped Ge2Sb2Te5 films studied by *in situ* resistance measurements," *Appl Phys Lett*, vol. 85, no. 15, pp. 3044–3046, Oct. 2004, doi: 10.1063/1.1805200.

[66] S. L. Firebaugh, K. F. Jensen, and M. A. Schmidt, "Investigation of high-temperature degradation of platinum thin films with an in situ resistance measurement apparatus," *Journal of Microelectromechanical Systems*, vol. 7, no. 1, pp. 128–135, Mar. 1998, doi: 10.1109/84.661395.

[67] M. Schuisky, J. W. Elam, and S. M. George, "*In situ* resistivity measurements during the atomic layer deposition of ZnO and W thin films," *Appl Phys Lett*, vol. 81, no. 1, pp. 180–182, Jul. 2002, doi: 10.1063/1.1490413.

[68] A. S. Ingason, F. Magnus, J. S. Agustsson, S. Olafsson, and J. T. Gudmundsson, "In-situ electrical characterization of ultrathin TiN films grown by reactive dc magnetron sputtering on SiO2," *Thin Solid Films*, vol. 517, no. 24, pp. 6731–6736, Oct. 2009, doi: 10.1016/J.TSF.2009.05.028.

[69] F. Magnus, A. S. Ingason, S. Olafsson, and J. T. Gudmundsson, "Growth and in-situ electrical characterization of ultrathin epitaxial TiN films on MgO," *Thin Solid Films*, vol. 519, no. 18, pp. 5861–5867, Jul. 2011, doi: 10.1016/J.TSF.2011.02.078.

[70] S. Li *et al.*, "Wafer-Scale and Deterministic Patterned Growth of Monolayer MoS2 via Vapor-Liquid-Solid Method," *Nanoscale*, vol. 11, no. 34, pp. 16122–16129, Jun. 2019, doi: 10.1039/C9NR04612G.

[71] H. Salami, A. Poissant, and R. A. Adomaitis, "Anomalously high alumina atomic layer deposition growth per cycle during trimethylaluminum under-dosing conditions," *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, vol. 35, no. 1, p. 01B101, Jan. 2017, doi: 10.1116/1.4963368.

[72] M. Reinke, Y. Kuzminykh, and P. Hoffmann, "Low Temperature Chemical Vapor Deposition Using Atomic Layer Deposition Chemistry," *Chemistry of Materials*, vol. 27, no. 5, pp. 1604–1611, Mar. 2015, doi: 10.1021/cm504216p.

[73] W. Lei, L. Henn-Lecordier, M. Anderle, G. W. Rubloff, M. Barozzi, and M. Bersani, "Real-time observation and optimization of tungsten atomic layer deposition process cycle," *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, vol. 24, no. 2, p. 780, Mar. 2006, doi: 10.1116/1.2184320.

[74] R. L. Z. Hoye, D. Muñoz-Rojas, K. P. Musselman, Y. Vaynzof, and J. L. MacManus-Driscoll, "Synthesis and Modeling of Uniform Complex Metal Oxides by Close-Proximity Atmospheric Pressure Chemical Vapor Deposition," *ACS Appl Mater Interfaces*, vol. 7, no. 20, pp. 10684–10694, May 2015, doi: 10.1021/am5073589.

[75] K. P. Musselman *et al.*, "Rapid open-air deposition of uniform, nanoscale, functional coatings on nanorod arrays," *Nanoscale Horiz*, vol. 2, no. 2, pp. 110–117, Feb. 2017, doi: 10.1039/C6NH00197A.

[76] S. M. George, "Atomic Layer Deposition: An Overview," *Chem Rev*, vol. 110, no. 1, pp. 111–131, Jan. 2010, doi: 10.1021/cr900056b.

[77] G. N. Parsons *et al.*, "Mechanisms and reactions during atomic layer deposition on polymers," *Coord Chem Rev*, vol. 257, no. 23–24, pp. 3323–3331, Dec. 2013, doi: 10.1016/J.CCR.2013.07.001.

[78] Zhuo Chen, K. Shum, T. Salagaj, Wei Zhang, and K. Strobl, "ZnO thin films synthesized by chemical vapor deposition," in *2010 IEEE Long Island Systems, Applications and Technology Conference*, IEEE, Jun. 2010, pp. 1–6. doi: 10.1109/LISAT.2010.5478331.

[79] H. K. Park, J. Jo, H. K. Hong, and J. Heo, "Influence of post-deposition annealing on the electrical properties of zinc oxide thin films," *Thin Solid Films*, vol. 573, pp. 22–26, Dec. 2014, doi: 10.1016/J.TSF.2014.10.100.

[80] A. Yamada, B. Sang, and M. Konagai, "Atomic layer deposition of ZnO transparent conducting oxides," *Appl Surf Sci*, vol. 112, pp. 216–222, Mar. 1997, doi: 10.1016/S0169-4332(96)01022-7.

[81] A. Dutta and S. Basu, "Modified CVD growth and characterization of ZnO thin films," *Mater Chem Phys*, vol. 34, no. 1, pp. 41–45, Apr. 1993, doi: 10.1016/0254-0584(93)90117-5.

[82] K. Ellmer, "Resistivity of polycrystalline zinc oxide films: current status and physical limit," *J Phys D Appl Phys*, vol. 34, no. 21, pp. 3097–3108, Nov. 2001, doi: 10.1088/0022-3727/34/21/301.

[83] T. K. Roy, D. Sanyal, D. Bhowmick, and A. Chakrabarti, "Temperature dependent resistivity study on zinc oxide and the role of defects," *Mater Sci Semicond Process*, vol. 16, no. 2, pp. 332–336, Apr. 2013, doi: 10.1016/J.MSSP.2012.09.018.

[84] E. Guziewicz *et al.*, "ALD grown zinc oxide with controllable electrical properties," *Semicond Sci Technol*, vol. 27, no. 7, p. 074011, Jun. 2012, doi: 10.1088/0268-1242/27/7/074011.

[85] T. A. Krajewski *et al.*, "Dominant shallow donors in zinc oxide layers obtained by low-temperature atomic layer deposition: Electrical and optical investigations," *Acta Mater*, vol. 65, pp. 69–75, Feb. 2014, doi: 10.1016/J.ACTAMAT.2013.11.054.

[86] G. Schleder, A. Padilha, C. Acosta, M. Costa, and A. Fazzio, "From DFT to Machine Learning: recent approaches to Materials Science–a review," *J. Phys. Materials*, May 2019, [Online]. Available: https://iopscience.iop.org/article/10.1088/2515-7639/ab084b/pdf

[87] C. Wen *et al.*, "Machine learning assisted design of high entropy alloys with desired property," *Acta Mater*, Mar. 2019, doi: 10.1016/J.ACTAMAT.2019.03.010.

[88] M. Wang, T. Wang, P. Cai, and X. Chen, "Nanomaterials Discovery and Design through Machine Learning," *Small Methods*, p. 1900025, Feb. 2019, doi: 10.1002/smtd.201900025.

[89] B. Meredig *et al.*, "Combinatorial screening for new materials in unconstrained composition space with machine learning," *Phys Rev B*, vol. 89, no. 9, p. 094104, Mar. 2014, doi: 10.1103/PhysRevB.89.094104.

[90] H. Li, Z. Zhang, Z.-Z. Zhao, H. Li, Z. Zhang, and Z.-Z. Zhao, "Data-Mining for Processes in Chemistry, Materials, and Engineering," *Processes*, vol. 7, no. 3, p. 151, Mar. 2019, doi: 10.3390/pr7030151.

[91] G. Estevez, N. Chen, C. He, and W. Zhu, "Lightweight Machine-Learning Model for Efficient Design of Graphene-Based Microwave Metasurfaces for Versatile Absorption Performance," *Nanomaterials 2023, Vol. 13, Page 329*, vol. 13, no. 2, p. 329, Jan. 2023, doi: 10.3390/NANO13020329.

[92] L.-K. Wan, Y.-X. Xue, J.-W. Jiang, and H. S. Park, "Machine learning accelerated search of the strongest graphene/h-BN interface with designed fracture properties," *J Appl Phys*, vol. 133, no. 2, p. 024302, Jan. 2023, doi: 10.1063/5.0131576.

[93] C. Rao and Y. Liu, "Three-dimensional convolutional neural network (3D-CNN) for heterogeneous material homogenization," *Comput Mater Sci*, vol. 184, p. 109850, Nov. 2020, doi: 10.1016/J.COMMATSCI.2020.109850.

[94] K. T. Schütt, F. Arbabzadah, S. Chmiela, K. R. Müller, and A. Tkatchenko, "Quantum-chemical insights from deep tensor neural networks," *Nat Commun*, vol. 8, p. 13890, Jan. 2017, doi: 10.1038/ncomms13890.

[95] I. Lobzenko, Y. Shiihara, H. Mori, and T. Tsuru, "Influence of group IV element on basic mechanical properties of BCC medium-entropy alloys using machine-learning potentials," *Comput Mater Sci*, vol. 219, p. 112010, Feb. 2023, doi: 10.1016/J.COMMATSCI.2023.112010.

[96] A. Yu-Tung Wang, M. Salah Mahmoud, M. Czasny, and A. Gurlo, "CrabNet for Explainable Deep Learning in Materials Science: Bridging the Gap Between Academia and Industry," *Integr Mater Manuf Innov*, vol. 11, pp. 41–56, Jan. 2022, doi: 10.1007/s40192-021-00247-y.

[97] N. Artrith and A. M. Kolpak, "Grand canonical molecular dynamics simulations of Cu–Au nanoalloys in thermal equilibrium using reactive ANN potentials," *Comput Mater Sci*, vol. 110, pp. 20–28, Dec. 2015, doi: 10.1016/J.COMMATSCI.2015.07.046.

[98] G. V. S. M. Carrera, L. C. Branco, J. Aires-de-Sousa, and C. A. M. Afonso, "Exploration of quantitative structure–property relationships (QSPR) for the design of new guanidinium ionic liquids," *Tetrahedron*, vol. 64, no. 9, pp. 2216–2224, Feb. 2008, doi: 10.1016/J.TET.2007.12.021.

[99] G. Hautier, C. Fischer, V. Ehrlacher, A. Jain, and G. Ceder, "Data Mined Ionic Substitutions for the Discovery of New Compounds," *Inorg Chem*, vol. 50, no. 2, pp. 656–663, Jan. 2011, doi: 10.1021/ic102031h.

[100] L. Zhang, Z. Chen, J. Su, and J. Li, "Data mining new energy materials from structure databases," *Renewable and Sustainable Energy Reviews*, vol. 107, pp. 554–567, Jun. 2019, doi: 10.1016/J.RSER.2019.03.036.

[101] Y. Mohammadi *et al.*, "Intelligent Machine Learning: Tailor-Making Macromolecules," *Polymers (Basel)*, vol. 11, no. 4, pp. 579–592, 2019, doi: 10.3390/polym11040579.

[102] B. Sanchez-Lengeling and A. Aspuru-Guzik, "Inverse molecular design using machine learning: Generative models for matter engineering.," *Science (1979)*, vol. 361, no. 6400, pp. 360–365, Jul. 2018, doi: 10.1126/science.aat2663.

[103] J. Su, G. Xiao, H. Zhang, and B. Li, "Microstructure reconstruction of porous copper foams based on deep convolutional generative adversarial networks with physical characteristics of materials," *Int J Data Sci Anal*, vol. 14, no. 1, pp. 89–97, Jun. 2022, doi: 10.1007/S41060-021-00308-7/FIGURES/13.

[104] R. Jabbar, R. Jabbar, and S. Kamoun, "Recent progress in generative adversarial networks applied to inversely designing inorganic materials: A brief review," *Comput Mater Sci*, vol. 213, p. 111612, Oct. 2022, doi: 10.1016/J.COMMATSCI.2022.111612.

[105] Y. Zhang and X. Xu, "Yttrium barium copper oxide superconducting transition temperature modelling through gaussian process regression," *Comput Mater Sci*, vol. 179, p. 109583, Jun. 2020, doi: 10.1016/J.COMMATSCI.2020.109583.

[106] V. L. Deringer, A. P. Bartók, N. Bernstein, D. M. Wilkins, M. Ceriotti, and G. Csányi, "Gaussian Process Regression for Materials and Molecules," *Chem Rev*, vol. 121, no. 16, pp. 10073–10141, Aug. 2021, doi: 10.1021/ACS.CHEMREV.1C00022/ASSET/IMAGES/MEDIUM/CR1C00022_M069.GIF.

[107] J. Gao, C. Wang, Z. Xu, J. Wang, S. Yan, and Z. Wang, "Gaussian process regression based remaining fatigue life prediction for metallic materials under two-step loading," *Int J Fatigue*, vol. 158, p. 106730, May 2022, doi: 10.1016/J.IJFATIGUE.2022.106730.

[108] G. Hautier, "Finding the needle in the haystack: Materials discovery and design through computational ab initio high-throughput screening," *Comput Mater Sci*, vol. 163, pp. 108–116, Jun. 2019, doi: 10.1016/J.COMMATSCI.2019.02.040.

[109] G. Hautier, C. C. Fischer, A. Jain, T. Mueller, and G. Ceder, "Finding Nature's Missing Ternary Oxide Compounds Using Machine Learning and Density Functional Theory," *Chemistry of Materials*, vol. 22, no. 12, pp. 3762–3767, Jun. 2010, doi: 10.1021/cm100795d.

[110] A. R. Singh, B. A. Rohr, J. A. Gauthier, and J. K. Nørskov, "Predicting Chemical Reaction Barriers with a Machine Learning Model," *Catal Letters*, pp. 1–8, Mar. 2019, doi: 10.1007/s10562-019-02705-x.

[111] P. V. Balachandran, "Machine learning guided design of functional materials with targeted properties," *Comput Mater Sci*, vol. 164, pp. 82–90, Jun. 2019, doi: 10.1016/J.COMMATSCI.2019.03.057.

[112] J. Syarif, M. B. Elbeltagy, and A. B. Nassif, "A machine learning framework for discovering high entropy alloys phase formation drivers," *Heliyon*, p. e12859, Jan. 2023, doi: 10.1016/J.HELIYON.2023.E12859.

[113] L. Huang and C. Ling, "Practicing deep learning in materials science: An evaluation for predicting the formation energies," *J Appl Phys*, vol. 128, no. 12, p. 124901, Sep. 2020, doi: 10.1063/5.0012411.

[114] F. Lu *et al.*, "Prediction of amorphous forming ability based on artificial neural network and convolutional neural network," *Comput Mater Sci*, vol. 210, p. 111464, Jul. 2022, doi: 10.1016/J.COMMATSCI.2022.111464.

[115] R. A. Hoyt, M. M. Montemore, I. Fampiou, W. Chen, G. Tritsaris, and E. Kaxiras, "Machine Learning Prediction of H Adsorption Energies on Ag Alloys," *J Chem Inf Model*, p. acs.jcim.8b00657, Mar. 2019, doi: 10.1021/acs.jcim.8b00657.

[116] J. Im, S. Lee, T.-W. Ko, H. W. Kim, Y. Hyon, and H. Chang, "Identifying Pb-free perovskites for solar cells by machine learning," *NPJ Comput Mater*, vol. 5, no. 1, p. 37, Dec. 2019, doi: 10.1038/s41524-019-0177-0.

[117] K. Zhou, Y. Hu, H. Pan, X. Gao, L. Wang, and L. Yao, "Porosity Prediction of Ceramic Matrix Composites Based on Random Forest," in *IOP Conference Series: Materials Science and Engineering*, Mar. 2020. doi: 10.1088/1757-899X/768/5/052115.

[118] S. A. Tawfik, M. Rashid, S. Gupta, S. P. Russo, T. R. Walsh, and S. Venkatesh, "Machine learning-based discovery of vibrationally stable materials," *npj Computational Materials 2023 9:1*, vol. 9, no. 1, pp. 1–6, Jan. 2023, doi: 10.1038/s41524-022-00943-z.

[119] C. L. Phillips and G. A. Voth, "Discovering crystals using shape matching and machine learning," *Soft Matter*, vol. 9, no. 35, p. 8552, Aug. 2013, doi: 10.1039/c3sm51449h.

[120] P. Raccuglia *et al.*, "Machine-learning-assisted materials discovery using failed experiments," *Nature*, vol. 533, no. 7601, pp. 73–76, May 2016, doi: 10.1038/nature17439.

[121] B. Lee, K. Yu, J. Jeon, and E. J. Choi, "Machine learning analysis of broadband optical reflectivity of semiconductor thin film," *Journal of the Korean Physical Society*, vol. 80, no. 4, pp. 347–351, Feb. 2022, doi: 10.1007/S40042-022-00436-8/FIGURES/4.

[122] E. Grau-Luque *et al.*, "Thickness evaluation of AlOx barrier layers for encapsulation of flexible PV modules in industrial environments by normal reflectance and machine learning," *Progress in Photovoltaics: Research and Applications*, vol. 30, no. 3, pp. 229–239, Mar. 2022, doi: 10.1002/PIP.3478.

[123] H. Tompkins, *A User's Guide to Ellipsometry*. Mesa, Arizona: Elsevier, 1993. doi: 10.1016/C2009-0-22336-1.

[124] A. Arunachalam, S. Novia Berriel, P. Banerjee, and K. Basu, "Machine Learning-enhanced Efficient Spectroscopic Ellipsometry Modeling," *ArXiv*, Jan. 2022, Accessed: Jan. 27, 2022. [Online]. Available: www.aaai.org

[125] T. Alcaire, D. Le Cunff, V. Gredy, and J. H. Tortai, "Spectroscopic ellipsometry imaging for process deviation detection via machine learning approach," *ASMC (Advanced Semiconductor Manufacturing Conference) Proceedings*, vol. 2020-August, Aug. 2020, doi: 10.1109/ASMC49169.2020.9185349.

[126] E. Fix and J. L. Hodges, "Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties," *Int Stat Rev*, vol. 57, no. 3, p. 238, Dec. 1989, doi: 10.2307/1403797.

[127] J. Liu, D. Zhang, D. Yu, M. Ren, and J. Xu, "Machine learning powered ellipsometry," *Light: Science & Applications 2021 10:1*, vol. 10, no. 1, pp. 1–7, Mar. 2021, doi: 10.1038/s41377-021-00482-0.

[128] A. L. Cauchy, *Œuvres complètes d'Augustin Cauchy, 1789-1857*. Paris: Gauthier-Villars et fils , 1882.

[129] G. E. Jellison and F. A. Modine, "Parameterization of the optical functions of amorphous materials in the interband region," *Appl Phys Lett*, vol. 69, no. 3, p. 371, Jun. 1998, doi: 10.1063/1.118064.

[130] J. Tauc, "Optical Properties and Electronic Structure of Amorphous Ge and Si," *Mater. Res. bull*, vol. 3, pp. 37–46, Nov. 1968.

[131] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," *Psychol Rev*, vol. 65, no. 6, pp. 386–408, 1958, doi: 10.1037/h0042519.

[132] M. Minsky and S. Papert, *Perceptrons: An Introduction to Computational Geometry - Marvin Minsky, Seymour A. Papert - Google Books*, 1st ed. Cambridge: MIT Press, 1969.

[133] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359–366, Jan. 1989, doi: 10.1016/0893-6080(89)90020-8.

[134] K.-I. Funahashi, "On the approximate realization of continuous mappings by neural networks," *Neural Networks*, vol. 2, no. 3, pp. 183–192, Jan. 1989, doi: 10.1016/0893-6080(89)90003-8.

[135] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, Oct. 1986, doi: 10.1038/323533a0.

[136] M. Claesen and B. De Moor, "Hyperparameter Search in Machine Learning," Feb. 2015, Accessed: Apr. 11, 2019. [Online]. Available: http://arxiv.org/abs/1502.02127

[137] J. Kiefer and J. Wolfowitz, "Stochastic Estimation of the Maximum of a Regression Function," *Ann. Math. Statist.*, vol. 23, no. 3, pp. 462–466, Sep. 1952, doi: 10.1214/aoms/1177705148.

[138] L. Bottou, F. E. Curtis, and J. Nocedal, "Optimization Methods for Large-Scale Machine Learning," *ArXiv*, Feb. 2018, doi: 10.1063/1.2965593.

[139] B. T. Polyak, "Some methods of speeding up the convergence of iteration methods," *USSR Computational Mathematics and Mathematical Physics*, vol. 4, no. 5, pp. 1–17, Jan. 1964, doi: 10.1016/0041-5553(64)90137-5.

[140] I. Sutskever, J. Martens, G. Dahl, and G. E. Hinton, "On the importance of initialization and momentum in deep learning," in *Proceedings of the 30th International Conference on Machine Learning*, Atlanta, 2013. doi: 10.1109/ICASSP.2013.6639346.

[141] M. Umehara, H. S. Stein, D. Guevarra, P. F. Newhouse, D. A. Boyd, and J. M. Gregoire, "Analyzing machine learning models to accelerate generation of fundamental materials insights," *NPJ Comput Mater*, vol. 5, no. 1, p. 34, Dec. 2019, doi: 10.1038/s41524-019-0172-5.

[142] H. Habibi Aghdam, E. Jahani Heravi, and D. Puig, "A practical approach for detection and classification of traffic signs using Convolutional Neural Networks," *Rob Auton Syst*, vol. 84, pp. 97–112, Oct. 2016, doi: 10.1016/J.ROBOT.2016.07.003.

[143] R. Maalej and M. Kherallah, "Convolutional Neural Network and BLSTM for Offline Arabic Handwriting Recognition," in *2018 International Arab Conference on Information Technology (ACIT)*, IEEE, Nov. 2018, pp. 1–6. doi: 10.1109/ACIT.2018.8672667.

[144] T. Liu, C. Rosenberg, and H. A. Rowley, "Clustering billions of images with large scale nearest neighbor search," in *Proceedings - IEEE Workshop on Applications of Computer Vision, WACV 2007*, Feb. 2007. doi: 10.1109/WACV.2007.18.

[145] D. Pradeepkumar and V. Ravi, "Forecasting financial time series volatility using Particle Swarm Optimization trained Quantile Regression Neural Network," *Appl Soft Comput*, vol. 58, pp. 35–52, Sep. 2017, doi: 10.1016/J.ASOC.2017.04.014.

[146] V. Nair and G. E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," in *Proceedings of the 27th International Conference on Machine Learning*, 2010. doi: 10.1.1.165.6419.

[147] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proceedings of the 30th International Conference on Machine Learning*, 2013, p. 6. doi: 10.1016/0010-0277(84)90022-2.

[148] G. Thimm and E. Fiesler, "High-order and multilayer perceptron initialization," *IEEE Trans Neural Netw*, vol. 8, no. 2, pp. 349–359, Mar. 1997, doi: 10.1109/72.557673.

[149] J. S. Bridle, "Training Stochastic Model Recognition Algorithms as Networks can Lead to Maximum Mutual Information Estimation of Parameters," *Adv Neural Inf Process Syst*, vol. 15, no. Ml, pp. 211–217, 1990.

[150] B. Leo, "Bagging predictors," *Mach Learn*, vol. 24, no. 2, pp. 123–140, Aug. 1996, [Online]. Available: http://dx.doi.org/10.1007/BF00058655

[151] R. E. Schapire, "The Strength of Weak Learnability," *Mach Learn*, vol. 5, pp. 197–227, Jun. 1990, doi: 10.1002/nbm.1810.

[152] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," *proc. OF THE IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998, [Online]. Available: http://ieeexplore.ieee.org/document/726791/#full-text-section

[153] M. Stone, "Cross-Validatory Choice and Assessment of Statistical Predictions," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 36, no. 2, pp. 111–133, Jan. 1974, doi: 10.1111/j.2517-6161.1974.tb00994.x.

[154] D. Srivastava and L. Bhambhu, "Data Classification Using Support Vector Machine," *J Theor Appl Inf Technol*, Feb. 2010.

[155] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?," Nov. 2014, Accessed: Apr. 14, 2019. [Online]. Available: http://arxiv.org/abs/1411.1792

[156] Z. Liu, C. T. Wu, and M. Koishi, "Transfer learning of deep material network for seamless structure–property predictions," *Comput Mech*, pp. 1–15, Apr. 2019, doi: 10.1007/s00466-019-01704-4.

[157] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*. Boston: The MIT Press, 2006.

[158] G. B. Goh, C. Siegel, A. Vishnu, N. Hodas, and N. Baker, "How Much Chemistry Does a Deep Neural Network Need to Know to Make Accurate Predictions?," *Proceedings - 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018*, vol. 2018-Janua, no. October 2017, pp. 1340–1349, 2018, doi: 10.1109/WACV.2018.00151.

[159] G. Goh, C. Siegel, N. Hodas, and A. Vishnu, "Using Rule-Based Labels for Weak Supervised Learning A ChemNet for Transferable Chemical Property Prediction," in *KDD London 2018*, Mar. 2018. doi: 10.475/123.

[160] S. Bishnoi, S. Singh, R. Ravinder, M. Bauchy, and N. N. Gosvami, "Predicting Youngs Modulus of Glasses with Sparse Datasets using Machine Learning," *ArXiv*.

[161] M. Picklum and M. Beetz, "MatCALO: Knowledge-enabled machine learning in materials science," *Comput Mater Sci*, vol. 163, pp. 50–62, Jun. 2019, doi: 10.1016/J.COMMATSCI.2019.03.005.

[162] Y. Marcus, "Solubilities of Buckminsterfullerene and Sulfur Hexafluoride in Various Solvents," *J. Phys. Chem. B*, vol. 101, no. 42, pp. 8617–8623, Oct. 1997, doi: 10.1021/JP970671S.

[163] M. A. Al-Khedher, C. Pezeshki, J. L. McHale, and F. J. Knorr, "Quality classification via Raman identification and SEM analysis of carbon nanotube bundles using artificial neural networks," *Nanotechnology*, vol. 18, no. 35, p. 355703, Sep. 2007, doi: 10.1088/0957-4484/18/35/355703.

[164] M. K. Sadan, H.-J. Ahn, G. S. Chauhan, and N. S. Reddy, "Quantitative estimation of poly(methyl methacrylate) nano-fiber membrane diameter by artificial neural networks," *Eur Polym J*, vol. 74, pp. 91–100, Jan. 2016, doi: 10.1016/J.EURPOLYMJ.2015.11.014.

[165] R. Hamzaoui, M. Cherigui, S. Guessasma, O. ElKedim, and N. Fenineche, "Artificial neural network methodology: Application to predict magnetic properties of nanocrystalline alloys," *Materials Science and Engineering: B*, vol. 163, no. 1, pp. 17–21, Jun. 2009, doi: 10.1016/J.MSEB.2009.04.015.

[166] M. A. Dobrovolskaia, D. R. Germolec, and J. L. Weaver, "Evaluation of nanoparticle immunotoxicity," *Nat Nanotechnol*, vol. 4, no. 7, pp. 411–414, Jul. 2009, doi: 10.1038/nnano.2009.175.

[167] G. Pyrgiotakis, O. E. Kundakcioglu, P. M. Pardalos, and B. M. Moudgil, "Raman spectroscopy and support vector machines for quick toxicological evaluation of titania nanoparticles," *Journal of Raman Spectroscopy*, vol. 42, no. 6, pp. 1222–1231, Jun. 2011, doi: 10.1002/jrs.2839.

[168] E. Oh *et al.*, "Meta-analysis of cellular toxicity for cadmium-containing quantum dots," *Nat Nanotechnol*, vol. 11, no. 5, pp. 479–486, May 2016, doi: 10.1038/nnano.2015.338.

[169] V. C. Epa, F. R. Burden, C. Tassa, R. Weissleder, S. Shaw, and D. A. Winkler, "Modeling Biological Activities of Nanoparticles," *Nano Lett*, vol. 12, no. 11, pp. 5808–5812, Nov. 2012, doi: 10.1021/nl303144k.

[170] M. Fernandez, H. Barron, and A. S. Barnard, "Artificial neural network analysis of the catalytic efficiency of platinum nanoparticles," *RSC Adv.*, vol. 7, no. 77, pp. 48962–48971, Oct. 2017, doi: 10.1039/C7RA06622H.

[171] R. Jinnouchi and R. Asahi, "Predicting Catalytic Activity of Nanoparticles by a DFT-Aided Machine-Learning Algorithm," *J Phys Chem Lett*, vol. 8, no. 17, pp. 4279–4283, Sep. 2017, doi: 10.1021/acs.jpclett.7b02010.

[172] N. Artrith and A. M. Kolpak, "Understanding the Composition and Activity of Electrocatalytic Nanoalloys in Aqueous Solvents: A Combination of DFT and Accurate Neural Network Potentials," *Nano Lett*, vol. 14, no. 5, pp. 2670–2676, May 2014, doi: 10.1021/nl5005674.

[173] M. Maghsoudi, M. Ghaedi, A. Zinali, A. M. Ghaedi, and M. H. Habibi, "Artificial neural network (ANN) method for modeling of sunset yellow dye adsorption using zinc oxide nanorods loaded on activated carbon: Kinetic and isotherm study," *Spectrochim Acta A Mol Biomol Spectrosc*, vol. 134, pp. 1–9, Jan. 2015, doi: 10.1016/J.SAA.2014.06.106.

[174] B. Sun, M. Fernandez, and A. S. Barnard, "Machine Learning for Silver Nanoparticle Electron Transfer Property Prediction," *J Chem Inf Model*, vol. 57, no. 10, pp. 2413–2423, Oct. 2017, doi: 10.1021/acs.jcim.7b00272.

[175] V. Simon, J. Gasteiger, and J. Zupan, "A Combined Application of Two Different Neural Network Types for the Prediction of Chemical Reactivity," *J Am Chem Soc*, vol. 115, no. 20, pp. 9148–9159, Oct. 1993, doi: 10.1021/ja00073a034.

[176] F. Amato, J. L. González-Hernández, and J. Havel, "Artificial neural networks combined with experimental design: A 'soft' approach for chemical kinetics," *Talanta*, vol. 93, pp. 72–78, May 2012, doi: 10.1016/j.talanta.2012.01.044.

[177] Z. Zhou, X. Li, and R. N. Zare, "Optimizing Chemical Reactions with Deep Reinforcement Learning," *ACS Cent Sci*, vol. 3, no. 12, pp. 1337–1344, Dec. 2017, doi: 10.1021/acscentsci.7b00492.

[178] J. Timoshenko, D. Lu, Y. Lin, and A. I. Frenkel, "Supervised Machine-Learning-Based Determination of Three-Dimensional Structure of Metallic Nanoparticles," *J Phys Chem Lett*, vol. 8, no. 20, pp. 5091–5098, Oct. 2017, doi: 10.1021/acs.jpclett.7b02364.

[179] C. Nyshadham *et al.*, "Machine-learned multi-system surrogate models for materials prediction," *NPJ Comput Mater*, vol. 5, no. 1, p. 51, Dec. 2019, doi: 10.1038/s41524-019-0189-9.

[180] Q. Zhu, A. Samanta, B. Li, R. E. Rudd, and T. Frolov, "Predicting phase behavior of grain boundaries with evolutionary search and machine learning," *Nat Commun*, vol. 9, no. 1, Feb. 2018, doi: 10.1038/s41467-018-02937-2.

[181] R. Y. Sato-Berrú, E. V. Basiuk (Golovataya-Dzhymbeeva), and J. M. Saniger, "Application of principal component analysis to discriminate the Raman spectra of functionalized multiwalled carbon nanotubes," *Journal of Raman Spectroscopy*, vol. 37, no. 11, pp. 1302–1306, Nov. 2006, doi: 10.1002/jrs.1561.

[182] X. Jiang *et al.*, "Surface-Enhanced Raman Nanoprobes with Embedded Standards for Quantitative Cholesterol Detection," *Small Methods*, vol. 2, no. 11, p. 1800182, Nov. 2018, doi: 10.1002/smtd.201800182.

[183] M. C. Swain and J. M. Cole, "ChemDataExtractor: A Toolkit for Automated Extraction of Chemical Information from the Scientific Literature," *J Chem Inf Model*, vol. 56, no. 10, pp. 1894–1904, Oct. 2016, doi: 10.1021/acs.jcim.6b00207.

[184] S. S. Kalantre *et al.*, "Machine learning techniques for state recognition and auto-tuning in quantum dots," *npj Quantum Inf*, vol. 5, no. 1, p. 6, Dec. 2019, doi: 10.1038/s41534-018-0118-7.

[185] Y. Huang *et al.*, "Band gap and band alignment prediction of nitride-based semiconductors using machine learning," *J Mater Chem C Mater*, Feb. 2019, doi: 10.1039/C8TC05554H.

[186] A. Majid, A. Khan, and T.-S. Choi, "Predicting lattice constant of complex cubic perovskites using computational intelligence," *Comput Mater Sci*, vol. 50, no. 6, pp. 1879–1888, Apr. 2011, doi: 10.1016/J.COMMATSCI.2011.01.035.

[187] A. Majid, A. Khan, G. Javed, and A. M. Mirza, "Lattice constant prediction of cubic and monoclinic perovskites using neural networks and support vector regression," *Comput Mater Sci*, vol. 50, no. 2, pp. 363–372, Dec. 2010, doi: 10.1016/J.COMMATSCI.2010.08.028.

[188] S. G. Javed, A. Khan, A. Majid, A. M. Mirza, and J. Bashir, "Lattice constant prediction of orthorhombic ABO3 perovskites using support vector machines," *Comput Mater Sci*, vol. 39, no. 3, pp. 627–634, May 2007, doi: 10.1016/J.COMMATSCI.2006.08.015.

[189] W. Huang, P. Martin, and H. L. Zhuang, "Machine-learning phase prediction of high-entropy alloys," *Acta Mater*, Mar. 2019, doi: 10.1016/J.ACTAMAT.2019.03.012.

[190] S. Liu *et al.*, "Convolutional neural networks for grazing incidence x-ray scattering patterns: thin film structure identification," *MRS Commun*, pp. 1–7, Mar. 2019, doi: 10.1557/mrc.2019.26.

[191] Z. Zhang, Y. Hong, B. Hou, Z. Zhang, M. Negahban, and J. Zhang, "Accelerated discoveries of mechanical properties of graphene using machine learning and high-throughput computation," *Carbon N Y*, Mar. 2019, doi: 10.1016/J.CARBON.2019.03.046.

[192] P. Hundi and R. Shahsavari, "Deep Learning to Speed up the Development of Structure–Property Relations For Hexagonal Boron Nitride and Graphene," *Small*, p. 1900656, Apr. 2019, doi: 10.1002/smll.201900656.

[193] E. J. Braham, J. Cho, K. M. Forlano, D. F. Watson, R. Arroyave, and S. Banerjee, "Machine-Learning-Directed Navigation of Synthetic Design Space: A Statistical Learning Approach to Controlling the Synthesis of Perovskite Halide Nanoplatelets in the Quantum-Confined Regime," *Chemistry of Materials*, p. acs.chemmater.9b00212, Apr. 2019, doi: 10.1021/acs.chemmater.9b00212.

[194] Z.-H. Shen *et al.*, "Phase-field modeling and machine learning of electric-thermal-mechanical breakdown of polymer-based dielectrics," *Nat Commun*, vol. 10, no. 1, p. 1843, Dec. 2019, doi: 10.1038/s41467-019-09874-8.

[195] N. Borodinov, S. Neumayer, S. V. Kalinin, O. S. Ovchinnikova, R. K. Vasudevan, and S. Jesse, "Deep neural networks for understanding noisy data applied to physical property extraction in scanning probe microscopy," *NPJ Comput Mater*, vol. 5, no. 1, p. 25, Dec. 2019, doi: 10.1038/s41524-019-0148-5.

[196] M. Ziatdinov *et al.*, "Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations," *ACS Nano*, vol. 11, no. 12, pp. 12742–12752, Dec. 2017, doi: 10.1021/acsnano.7b07504.

[197] M. P. Nikiforov *et al.*, "Functional recognition imaging using artificial neural networks: applications to rapid cellular identification via broadband electromechanical response," *Nanotechnology*, vol. 20, no. 40, p. 405708, Oct. 2009, doi: 10.1088/0957-4484/20/40/405708.

[198] T. Chen, Y. Zhou, and M. Rafailovich, "Application of Machine Learning in Perovskite Solar Cell Crystal Size Distribution Analysis," *MRS Adv*, pp. 1–8, Feb. 2019, doi: 10.1557/adv.2019.145.

[199] M. Sluydts, M. Larmuseau, J. Lauwaert, and S. Cottenier, "The road to accuracy : machine-learning-accelerated silicon ab initio simulations," in *The 8th Forum on the Science and Technology of Silicon Materials 2018*, 2018.

[200] R. Jinnouchi, J. Lahnsteiner, F. Karsai, G. Kresse, and M. Bokdam, "Phase transitions of hybrid perovskites simulated by machine-learning force fields trained on-the-fly with Bayesian inference," *ArXiv*, pp. 1–5, Jun. 2019.

[201] V. H. Nguyen, U. Gottlieb, A. Valla, D. Muñoz, D. Bellet, and D. Muñoz-Rojas, "Electron tunnelling through grain boundaries in transparent conductive oxides and implications for electrical conductivity: the case of ZnO:Al thin films," *Mater Horiz*, vol. 5, no. 4, pp. 715–726, Jul. 2018, doi: 10.1039/C8MH00402A.

[202] J. T-Thienprasert *et al.*, "Compensation in Al-doped ZnO by Al-related acceptor complexes: Synchrotron X-ray absorption spectroscopy and theory," *Phys Rev Lett*, vol. 110, no. 5, p. 055502, Jan. 2013, doi: 10.1103/PHYSREVLETT.110.055502/FIGURES/3/MEDIUM.

[203] H. Kwak *et al.*, "Angstrom-accuracy multilayer thickness determination using optical metrology and machine learning," in *SPIE Proceedings Vol. 11782*, SPIE, Jun. 2021, pp. 178–182. doi: 10.1117/12.2592216.

[204] A. Arunachalam *et al.*, "Machine learning approach to thickness prediction from in situ spectroscopic ellipsometry data for atomic layer deposition processes," *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, vol. 40, no. 1, p. 012405, Dec. 2021, doi: 10.1116/6.0001482.

[205] A. S. Yersak, Y. C. Lee, J. A. Spencer, and M. D. Groner, " Atmospheric pressure spatial atomic layer deposition web coating with in situ monitoring of film thickness ," *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, vol. 32, no. 1, p. 01A130, Dec. 2014, doi: 10.1116/1.4850176.

[206] D. Muñoz-Rojas, V. Huong Nguyen, C. Masse de la Huerta, C. Jiménez, and D. Bellet, "Spatial Atomic Layer Deposition," in *Chemical Vapor Deposition for Nanotechnology*, IntechOpen, 2019. doi: 10.5772/intechopen.82439.

[207] L. Hoffmann *et al.*, "Atmospheric pressure plasma enhanced spatial atomic layer deposition of SnO_x as conductive gas diffusion barrier," *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, vol. 36, no. 1, p. 01A112, Dec. 2018, doi: 10.1116/1.5006781.

[208] L. Hoffmann *et al.*, "Spatial Atmospheric Pressure Atomic Layer Deposition of Tin Oxide as an Impermeable Electron Extraction Layer for Perovskite Solar Cells with Enhanced Thermal Stability," *ACS Appl Mater Interfaces*, vol. 10, no. 6, pp. 6006–6013, Jan. 2018, doi: 10.1021/acsami.7b17701.

[209] Y. Zhang and X. Xu, "Machine learning optical band gaps of doped-ZnO films," *Optik (Stuttg)*, vol. 217, p. 164808, Sep. 2020, doi: 10.1016/J.IJLEO.2020.164808.

[210] M. Sajjad, I. Ullah, M. I. Khan, J. Khan, M. Y. Khan, and M. T. Qureshi, "Structural and optical properties of pure and copper doped zinc oxide nanoparticles," *Results Phys*, vol. 9, pp. 1301–1309, Jun. 2018, doi: 10.1016/J.RINP.2018.04.010.

[211] P. R. Regonia, C. M. Pelicano, R. Tani, A. Ishizumi, H. Yanagi, and K. Ikeda, "Predicting the band gap of ZnO quantum dots via supervised machine learning models," *Optik (Stuttg)*, vol. 207, p. 164469, Apr. 2020, doi: 10.1016/J.IJLEO.2020.164469.

[212] V. Srikant and D. R. Clarke, "On the optical band gap of zinc oxide," *J Appl Phys*, vol. 83, no. 10, pp. 5447–5451, May 1998, doi: 10.1063/1.367375.

[213] S. Muhammad, A. T. Nomaan, M. I. Idris, and M. Rashid, "Structural, optical and electrical investigation of low-temperature processed zinc oxide quantum dots based thin films using precipitation-spin coating on flexible substrates," *Physica B Condens Matter*, vol. 635, p. 413806, Jun. 2022, doi: 10.1016/J.PHYSB.2022.413806.

[214] M. Steglich, A. Bingel, G. Jia, and F. Falk, "Atomic layer deposited ZnO:Al for nanostructured silicon heterojunction solar cells," *Solar Energy Materials and Solar Cells*, vol. 103, pp. 62–68, Aug. 2012, doi: 10.1016/J.SOLMAT.2012.04.004.

[215] B. A. Anandh, R. Sakthivel, A. S. Ganesh, S. Subramani, and A. T. Rajamanickam, "ANALYZING ANTIMICROBIAL ACTIVITY OF ALUMINIUM DOPED ZnO THIN FILMS," *Rasayan Journal of Chemistry*, vol. 15, no. 1, pp. 387–394, Jan. 2022, doi: 10.31788/RJC.2022.1516626.

[216] H. Zhou, H. Wang, K. Zheng, Z. Gu, Z. Wu, and X. Tian, "Aluminum-doped zinc oxide nanoparticles with tunable near-infrared absorption/reflectance by a simple solvothermal process," *RSC Adv*, vol. 4, no. 81, pp. 42758–42763, Sep. 2014, doi: 10.1039/c4ra08196j.

[217] M. J. Akhtar, H. A. Alhadlaq, A. Alshamsan, M. A. Majeed Khan, and M. Ahamed, "Aluminum doping tunes band gap energy level as well as oxidative stress-mediated cytotoxicity of ZnO nanoparticles in MCF-7 cells," *Scientific Reports 2015 5:1*, vol. 5, no. 1, pp. 1–16, Sep. 2015, doi: 10.1038/srep13876.

[218] J. M. Khoshman and M. E. Kordesch, "Optical constants and band edge of amorphous zinc oxide thin films," *Thin Solid Films*, vol. 515, no. 18, pp. 7393–7399, Jun. 2007, doi: 10.1016/J.TSF.2007.03.055.

[219] J. Cañas *et al.*, "Determination of alumina bandgap and dielectric functions of diamond MOS by STEM-VEELS," *Appl Surf Sci*, vol. 461, pp. 93–97, Dec. 2018, doi: 10.1016/J.APSUSC.2018.06.163.

[220] S. Canulescu *et al.*, "Band gap tuning of amorphous Al oxides by Zr alloying," *Appl Phys Lett*, vol. 109, no. 9, p. 091902, Aug. 2016, doi: 10.1063/1.4961941.

[221] M. Reza Gaeeni, M. Sasani Ghamsari, A. Majd Abadi, M. Hossein Majles Ara, W. Han, and H.-H. Park, "Al 2 O 3 Colloidal Nanocrystals with Strong UV Emission," *Journal of the American Ceramic Society*, vol. 98, no. 6, pp. 1818–1822, Mar. 2015, doi: 10.1111/jace.13546.

[222] J. P. Colonna *et al.*, "Study of parasitic trapping in alumina used as blocking oxide for nonvolatile memories," *Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena*, vol. 29, no. 1, p. 01AE02, Jan. 2011, doi: 10.1116/1.3535552.

[223] S. S. Parhizgar, S. Taheriniya, and J. Beheshtian, "Experimental and Theoretical Study of Porous Al2O3," *Transactions of the Indian Institute of Metals*, vol. 74, no. 2, pp. 381–386, Feb. 2021, doi: 10.1007/S12666-020-02092-7/FIGURES/5.

[224] S.-D. Mo and W. Y. Ching, "Electronic and optical properties of Theta-Al 2 O 3 and comparison to alpha-Al 2 O 3," Jun. 1998.

[225] H. Dondapati, K. Santiago, and A. K. Pradhan, "Influence of growth temperature on electrical, optical, and plasmonic properties of aluminum:zinc oxide films grown by radio frequency magnetron sputtering," *J Appl Phys*, vol. 114, no. 14, Oct. 2013, doi: 10.1063/1.4824751.

[226] O. Reshef, I. De Leon, M. Z. Alam, and R. W. Boyd, "Nonlinear optical effects in epsilon-near-zero media," *Nature Reviews Materials 2019 4:8*, vol. 4, no. 8, pp. 535–551, Jun. 2019, doi: 10.1038/s41578-019-0120-5.

[227] M. Arif *et al.*, "A structural, morphological, linear, and nonlinear optical spectroscopic studies of nanostructured Al-doped ZnO thin films: An effect of Al concentrations," *J Mater Res*, pp. 1–9, Feb. 2019, doi: 10.1557/jmr.2018.506.

[228] B. Sarma, D. Barman, and B. K. Sarma, "AZO (Al:ZnO) thin films with high figure of merit as stable indium free transparent conducting oxide," *Appl Surf Sci*, vol. 479, pp. 786–795, Jun. 2019, doi: 10.1016/J.APSUSC.2019.02.146.

[229] K. T. Chee, V. V. Truong, and F. E. Girouard, "Interband transitions in aggregated copper films," *Applied Optics, Vol. 20, Issue 3, pp. 404-406*, vol. 20, no. 3, pp. 404–406, Feb. 1981, doi: 10.1364/AO.20.000404.

[230] A. R. Ansari, A. H. Hammad, M. S. Abdel-wahab, M. Shariq, and M. Imran, "Structural, optical and photoluminescence investigations of nanocrystalline CuO thin films at different microwave powers," *Opt Quantum Electron*, vol. 52, no. 10, p. 426, Oct. 2020, doi: 10.1007/s11082-020-02535-x.

[231] N. R. Dhineshbabu, • V Rajendran, • N Nithyavathy, and • R Vetumperumal, "Study of structural and optical properties of cupric oxide nanoparticles," *Appl Nanosci*, vol. 6, pp. 933–939, Sep. 2016, doi: 10.1007/s13204-015-0499-2.

[232] A. F. Qasrawi and A. A. Hamamdah, "Thickness effects on the dielectric dispersion and optical conductivity parameters of CuO thin films," *Microw Opt Technol Lett*, vol. 62, no. 4, pp. 1453–1458, Apr. 2020, doi: 10.1002/MOP.32192.

[233] V. Figueiredo *et al.*, "Effect of post-annealing on the properties of copper oxide thin films obtained from the oxidation of evaporated metallic copper," *Appl Surf Sci*, vol. 254, no. 13, pp. 3949–3954, Apr. 2008, doi: 10.1016/J.APSUSC.2007.12.019.

[234] C. Ravichandiran *et al.*, "In-depth study on structural, optical, photoluminescence and electrical properties of electrodeposited Cu 2 O thin films for optoelectronics: An effect of solution pH," *Microelectron Eng*, vol. 210, pp. 27–34, Apr. 2019, doi: 10.1016/j.mee.2019.03.013.

[235] H. Solache-Carranco *et al.*, "Photoluminescence and X-ray diffraction studies on Cu2O," *J Lumin*, vol. 129, no. 12, pp. 1483–1487, Dec. 2009, doi: 10.1016/j.jlumin.2009.02.033.

[236] G. Martínez-Saucedo, C. G. Torres-Castanedo, S. Arias-Cerón, R. Castanedo-Pérez, G. Torres-Delgado, and O. Zelaya-Ángel, "Photoluminescence of Cu2O nanostructured in stressed thin films induced by temperature," *J Lumin*, vol. 215, p. 116642, Nov. 2019, doi: 10.1016/j.jlumin.2019.116642.

[237] A. Sekkat *et al.*, "Chemical deposition of Cu2O films with ultra-low resistivity: correlation with the defect landscape," *Nature Communications 2022 13:1*, vol. 13, no. 1, pp. 1–11, Sep. 2022, doi: 10.1038/s41467-022-32943-4.

[238] A. S. Averyushkin *et al.*, "Stimulated low frequency Raman scattering in cupric oxide nanoparticles water suspension," *Opt Commun*, vol. 389, pp. 51–53, Apr. 2017, doi: 10.1016/j.optcom.2016.12.016.

[239] P. J. M. Isherwood, "Copper zinc oxide: Investigation into a p-type mixed metal oxide system," *Vacuum*, vol. 139, pp. 173–177, May 2017, doi: 10.1016/J.VACUUM.2016.09.026.

[240] L. Caspani *et al.*, "Enhanced Nonlinear Refractive Index in ϵ-Near-Zero Materials," *Phys. Rev. Lett.*, vol. 116, no. 23, Jun. 2016, doi: 10.1103/PhysRevLett.116.233901.

[241] G. V Naik, V. M. Shalaev, A. Boltasseva, G. V Naik, V. M. Shalaev, and A. Boltasseva, "Alternative Plasmonic Materials: Beyond Gold and Silver," *Advanced Materials*, vol. 25, no. 24, pp. 3264–3294, May 2013, doi: 10.1002/adma.201205076.

[242] C. T. Riley *et al.*, "High-Quality, Ultraconformal Aluminum-Doped Zinc Oxide Nanoplasmonic and Hyperbolic Metamaterials," *Small*, vol. 12, no. 7, pp. 892–901, Feb. 2016, doi: 10.1002/SMLL.201501797.

[243] O. Stenzel, S. Wilbrandt, S. Stempfhuber, D. Gäbler, and S. J. Wolleb, "Spectrophotometric characterization of thin copper and gold films prepared by electron beam evaporation: Thickness dependence of the drude damping parameter," *Coatings*, vol. 9, no. 3, Mar. 2019, doi: 10.3390/COATINGS9030181.

[244] M. A. Ordal, R. J. Bell, R. W. Alexander, L. L. Long, and M. R. Querry, "Optical properties of fourteen metals in the infrared and far infrared: Al, Co, Cu, Au, Fe, Pb, Mo, Ni, Pd, Pt, Ag, Ti, V, and W," *Appl Opt*, vol. 24, no. 24, p. 4493, Dec. 1985, doi: 10.1364/AO.24.004493.

[245] E. J. Zeman and G. C. Schatz, "An accurate electromagnetic theory study of surface enhancement factors for Ag, Au, Cu, Li, Na, Al, Ga, In, Zn, and Cd," *Journal of Physical Chemistry*, vol. 91, no. 3, pp. 634–643, Dec. 1987, doi: 10.1021/J100287A028/ASSET/J100287A028.FP.PNG_V03.

[246] M. García-Méndez, Á. Bedoya-Calle, R. R. Segura, and V. Coello, "Investigation of the annealing effects on the structural and optoelectronic properties of RF-sputtered ZnO films studied by the Drude–Lorentz model," *Appl Phys A Mater Sci Process*, vol. 120, no. 4, pp. 1375–1382, Sep. 2015, doi: 10.1007/S00339-015-9318-6/FIGURES/6.

[247] B. A. Gizhevskii, Y. P. Sukhorukov, N. N. Loshkareva, A. S. Moskvin, E. V Zenkov, and E. A. Kozlov, "Surface plasmon resonances in the absorption spectra of nanocrystalline cupric oxide," Nov. 2003.

[248] J. Zou and L. Schiebinger, "AI can be sexist and racist -- it's time to make it fair," *Nature*, vol. 559, no. 7714, pp. 324–327, Jul. 2018, doi: 10.1038/D41586-018-05707-8.

[249] M. T. Hörantner *et al.*, "The Potential of Multijunction Perovskite Solar Cells," *ACS Energy Lett*, vol. 2, no. 10, pp. 2506–2513, Oct. 2017, doi: 10.1021/ACSENERGYLETT.7B00647/SUPPL_FILE/NZ7B00647_SI_002.TXT.

[250] P. Yang *et al.*, "The investigation of CsPb(I1−xBrx)3/crystalline silicon two- and four-terminal tandem solar cells," *Solar Energy*, vol. 216, pp. 145–150, Mar. 2021, doi: 10.1016/J.SOLENER.2021.01.041.

[251] M. I. ; Asghar *et al.*, "Device stability of perovskite solar cells - A review," *Renewable and Sustainable Energy Reviews*, vol. 77, pp. 131–146, Sep. 2017, doi: 10.1016/j.rser.2017.04.003.

[252] K. Ibrahim *et al.*, "Humidity-resistant perovskite solar cells via the incorporation of halogenated graphene particles," *Solar Energy*, vol. 224, pp. 787–797, Aug. 2021, doi: 10.1016/J.SOLENER.2021.06.016.

[253] D. Yu, Y. Q. Yang, Z. Chen, Y. Tao, and Y. F. Liu, "Recent progress on thin-film encapsulation technologies for organic electronic devices," *Opt Commun*, vol. 362, pp. 43–49, Mar. 2016, doi: 10.1016/J.OPTCOM.2015.08.021.

[254] K. O. Brinkmann *et al.*, "Suppressed decomposition of organometal halide perovskites by impermeable electron-extraction layers in inverted solar cells," *Nature Communications 2017 8:1*, vol. 8, no. 1, pp. 1–9, Jan. 2017, doi: 10.1038/ncomms13938.

[255] J.-H. Woo *et al.*, "Amorphous Alumina Film Robust under Cyclic Deformation: a Highly Impermeable and a Highly Flexible Encapsulation Material," *Cite This: ACS Appl. Mater. Interfaces*, vol. 13, p. 46901, Sep. 2021, doi: 10.1021/acsami.1c15261.

[256] N. Titan Putri Hartono *et al.*, "How machine learning can help select capping layers to suppress perovskite degradation," *Nat Commun*, vol. 11, Aug. 2020, doi: 10.1038/s41467-020-17945-4.

[257] I. J. Goodfellow *et al.*, "Generative Adversarial Nets," *Adv Neural Inf Process Syst*, pp. 2672–2680, 2014, Accessed: Mar. 27, 2023. [Online]. Available: http://www.github.com/goodfeli/adversarial

[258] A. Makhzani, J. Shlens, N. Jaitly, G. Brain, I. G. Openai, and B. Frey, "Adversarial Autoencoders," Nov. 2015, Accessed: Mar. 27, 2023. [Online]. Available: https://arxiv.org/abs/1511.05644v2

[259] Z. Yang, X. Li, L. C. Brinson, A. N. Choudhary, W. Chen, and A. Agrawal, "Microstructural materials design via deep adversarial learning methodology," *Journal of Mechanical Design, Transactions of the ASME*, vol. 140, no. 11, p. 111416, Nov. 2018, doi: 10.1115/1.4041371.

[260] L. Caspani *et al.*, "Enhanced Nonlinear Refractive Index in ε-Near-Zero Materials," *Phys Rev Lett*, vol. 116, no. 23, Jun. 2016, doi: 10.1103/PHYSREVLETT.116.233901.

[261] N. Kinsey, C. DeVault, J. Kim, M. Ferrera, V. M. Shalaev, and A. Boltasseva, "Epsilon-Near-Zero Al-Doped ZnO for Ultrafast Switching at Telecom Wavelengths: Outpacing the Traditional Amplitude-Bandwidth Trade-Off," *Optica*, vol. 2, no. 7, p. 616, Mar. 2015, doi: 10.1364/OPTICA.2.000616.

[262] C. T. Riley *et al.*, "High-Quality, Ultraconformal Aluminum-Doped Zinc Oxide Nanoplasmonic and Hyperbolic Metamaterials," *Small*, vol. 12, no. 7, pp. 892–901, Feb. 2016, doi: 10.1002/SMLL.201501797.