Machine learning-assisted nanoscale photoelectrical sensing

Ziyan Zhu
Stanford Institute for Materials and Energy Sciences,
SLAC National Accelerator Laboratory,
Menlo Park, CA 94025
ziyanzhu@stanford.edu

Zhurun Ji
Department of Physics and Applied Physics,
Stanford University,
Stanford, CA 94305, USA
Gaballe Laboratory of Advanced Materials,
Stanford University,
Stanford, CA 94305, USA
zhurun@stanford.edu

Houssam Yassin
Department of Applied Mathematics,
University of Colorado, Boulder, CO 80309, USA
houssam.yassin@colorado.edu

Zhi-Xun Shen
Stanford Institute for Materials and Energy Sciences,
SLAC National Accelerator Laboratory,
Menlo Park, CA 94025
Department of Physics and Applied Physics,
Stanford University,
Stanford, CA 94305, USA
zxshen@stanford.edu

Thomas P. Devereaux
Stanford Institute for Materials and Energy Sciences,
SLAC National Accelerator Laboratory,
Menlo Park, CA 94025
Department of Materials Science and Engineering,
Stanford University,
CA 94305, USA
tpd@stanford.edu

Abstract

The ability to non-invasively measure local conductivity and permittivity at the nanoscale is of fundamental importance in unraveling the physics of quantum systems. One approach is Microwave Impedance Microscopy (MIM), a scanning probe technique operating at microwave frequencies. However, the resulting

Machine Learning and the Physical Sciences Workshop, NeurIPS 2023.

large datasets and vast parameter space make obtaining a mapping between MIM measurements and local microscopic properties challenging. Here, we overcome this challenge by using machine learning to reconstruct the local properties while incorporating physical priors. The synergy between MIM and ML allows for the quantitative predictions of complex interactions between excitons, charge carriers, and the dielectric environment. This approach provides profound insights into the fundamental physics of excitons in two-dimensional materials.

1 Introduction

The quest to unlock the properties of quantum materials at the nanoscale has led to the development of cutting-edge techniques. Among these, Microwave Impedance Microscopy (MIM) stands out as a powerful scanning probe method that has the capability to probe a sample's local properties with sub-100 nm spatial resolution, with applications ranging from semiconducting devices to biological specimens [1]. Current research aims to expand MIM's capabilities for materials discovery and development by realizing high-speed imaging, multimodal imaging and improved spatial resolution. The improved capabilities result in large datasets that call for new algorithms to extract meaningful physical insights from them.

In this work, we develop a machine learning-based approach to enhance the MIM capabilities. In particular, we focus on understanding local optoelectric and dielectric properties in monolayer transition metal dichalcogenides (TMDs), known for robust excitonic effects due to high binding energies, direct bandgap, and strong light-matter interactions. The measurements were done using advanced exciton-resonant microwave impedance microscopy (ER-MIM) that is highly sensitive to exciton photoconductivity and dielectric properties [2]. However, the data obtained from ER-MIM is massive and multidimensional, making manual analysis time-consuming and error-prone. To enhance the efficacy of the analysis, we integrate ER-MIM with ML algorithms to automate the analysis process, and harness their combined potential to quantitatively investigate the complex interactions among excitons, charge carriers, and the dielectric environment. This innovative synergy opens up new avenues for understanding complex quantum phenomena.

Main Contributions: We employ ML to predict optoelectric properties and dielectric characteristics of monolayer TMDs based on ER-MIM measurements. Our approach predicts the probability distributions of conductivity and dielectric properties, integrating physics knowledge into measurements. This integration of ML with experimental techniques enables precise material analysis at the nanoscale.

Related Work: ML has improved the analysis of scanning probe methods like scanning tunneling microscopy (STM) and atomic force microscopy (AFM), enhancing speed, noise reduction, automation, and property extraction [3, 4, 5, 6]. Scattering-type scanning near-field optical microscopy (s-SNOM) also benefits from ML, aiding denoising and phase classification [7]. However, in the intersection of optics and microwave impedance microscopy, the application of ML and AI is still emerging, an area we address in this work.

2 Deep neural network for local property predictions

MIM operates by positioning a microwave-driven metal tip near a sample's surface. The sample modifies the electric potential around the tip, which causes a portion of the electrical signal to reflect back along the microwave line, providing information about the sample's local properties and its dielectric environment. Figure 1 shows typical response curves of MIM as a function of conductivity. The sample is encapsulated by a 30 nm thick hexagonal Boron Nitride (hBN) substrate, and the substrate's permittivity, denoted as ϵ, influences the measurement. Change in the sample topography affects the tip-sample distance, denoted as h, with smaller values of h resulting in stronger signals.

Our goal is to predict the local environment (σ, ϵ, h) for a given set of MIM real and imaginary parts, $\mathrm{Re}(\Delta\mathrm{MIM})$ and $\mathrm{Im}(\Delta\mathrm{MIM})$. However, this mapping is not bijective, as multiple sets of (σ, ϵ, h) can correspond to the same ($\mathrm{Re}(\Delta\mathrm{MIM})$, $\mathrm{Im}(\Delta\mathrm{MIM})$). Therefore, instead of predicting a fixed (σ, ϵ, h), we assign a probability to each set of (σ, ϵ, h) for a given measurement. Figure 1 shows the network architecture used for predicting local properties. The network takes input values of

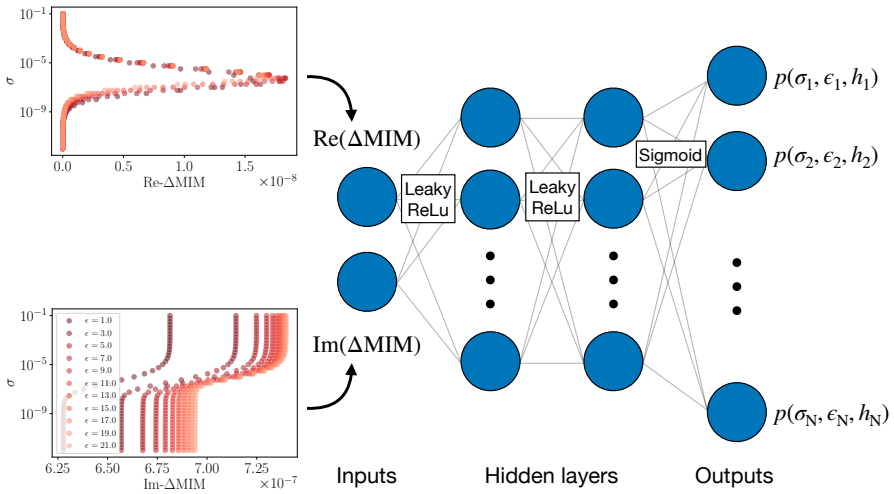

Figure 1: Architecture for predicting the probability distribution of local microscopic properties. Insets show the conductivity σ as a function of different values of the real and imaginary parts of the MIM for a fixed tip-sample height $h = 10$ nm, with each curve corresponding to a different substrate permittivity. The real and imaginary parts are used as inputs to the network. The outputs represent the probability for each given set of $(\sigma_j, \epsilon_j, h_j)$.

$\mathrm{Re}(\Delta\mathrm{MIM})$ and $\mathrm{Im}(\Delta\mathrm{MIM})$, either from experiments or simulations, and predicts the probability distribution for discretized (σ, ϵ, h) values.

For training the network, we start by simulating $\mathrm{Re}(\Delta\mathrm{MIM})$ and $\mathrm{Im}(\Delta\mathrm{MIM})$ for various sets of (σ, ϵ, h) (see Fig. 1 insets). We then interpolate this simulation data across a physically relevant range of (σ, ϵ, h) values: σ from 1×10^{-11} to 6×10^{-3} S/sq, ϵ between 3.2 and 3.9 [8], and h from 6 to 18 nm, with even spacing in each dimension. This amounts to a total of 64,000 sets of data. We further divide the data into bins: 32 bins for σ, 8 bins for ϵ, and 15 bins for h, resulting in 3,840 bins in total. Each data point is then encoded into a 3,840-dimensional one-hot encoding vector based on its bin membership. During training, we minimize the Binary Cross Entropy loss, transforming the prediction into a classification problem, where the network output is a 3,840-dimensional vector representing the probability for each bin.

The network architecture consists of 2 hidden layers with 100 nodes each. We use the Adam optimizer with a constant learning rate of 0.01 and set β_1 and β_2 to 0.999 and 0.9999, respectively. We train until the variance of the loss function over the last 200 steps falls below 1×10^{-5}. Data is randomly split into 80% for training, 10% for validation, and 10% for testing.

3 Results

Evaluation on the test data: We first assess the trained network's performance on synthetic data with noise added. Figure 2(a)-(b) shows simulated $\mathrm{Im}(\Delta\mathrm{MIM})$ and $\mathrm{Re}(\Delta\mathrm{MIM})$ at varying conductivity, while holding ϵ and h constant. We average over 100 samples drawn from the predicted probability distribution for each input set. Given the multivalued nature of the real and imaginary parts of MIM in ϵ and h, we condition the sampling of the initial sample at the first σ value around $\epsilon_0 = 3.5$ and $h_0 = 12$ nm. To achieve this, we calculate the distance, D, of each bin from (ϵ_0, h_0) as $D_j = \sqrt{(\epsilon_j - \epsilon_0)^2 + (h_j - h_0)^2}$ and weight the predicted probability distribution with $\exp(-\alpha D_j)$, where α denotes the constraint strength (with $\alpha = 2$ in this case). Subsequent samples are conditioned on the (ϵ, h) values of the preceding sample. Figure 2 (c)-(e) illustrates the neural network-predicted ϵ, σ, and h as functions of the ground truth σ. The predictions align with the ground truth (red dashed line) within error bars in all cases, indicating our model's accurate prediction capability for (σ, ϵ, h) in MIM measurements despite the noise.

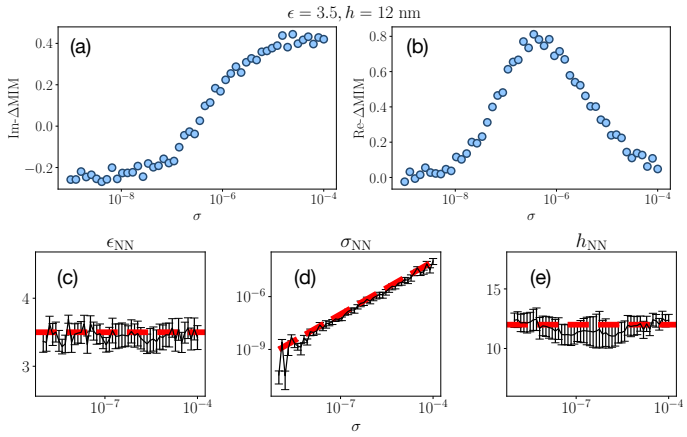

Figure 2: (a)-(b) Synthetic data obtained from interpolating simulations with random noise of amplitude 0.04, while varying conductivity σ and keeping permittivity and sample-tip distance fixed at $\epsilon = 3.5$ and $h = 12\,\mathrm{nm}$, respectively. (c)-(e) Predicted ϵ, σ, and h as functions of the ground truth conductivity, averaged over 100 samples. The red dashed lines represent the ground truth.

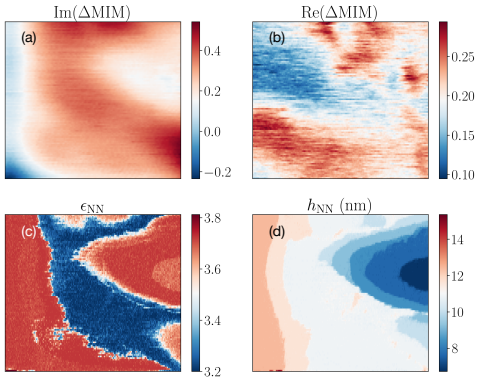

Figure 3: (a)-(b) Imaginary and real parts of MIM at $V_g = -2.25\,\mathrm{V}$ and wavelength $\lambda = 745\,\mathrm{nm}$. (c)-(d) predicted hBN permittivity (c) and height-sample distance (d).

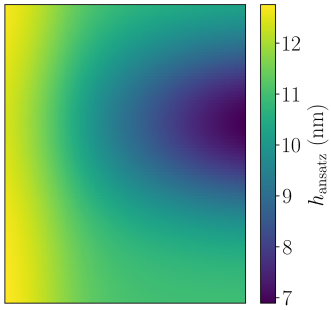

Figure 4: Height distribution used as an initial guess for Fig. 3.

Predictions from Experimental Data: We now use the trained model to predict the local environment and properties of a $MoSe_2$ device under light coupling based on experimental measurements. A graphite gate covers the right side of the measurement region, forming a p-n junction. By varying gate voltages on the right region and the optical wavelength, we measure the change in the ER-MIM response. Here, due to the coupling to the light, we measure the photoconductivity instead of the static electrical conductivity. A typical ER-MIM scan is shown in Fig. 3(a)-(b) for a $1\mu\mathrm{m}\times1\mu\mathrm{m}$ region. We take one set of scans for predicting ϵ and h and keep them fixed. This is because ϵ and h are determined by the substrate and the sample topography respectively and therefore are not expected to change from scan to scan. While we do not have access to the precise sample topography, based on the AFM image that reveals a bubble on the right and the existence of an edge on the left, we condition the sample on an initial guess height distribution shown in Figure 4 based on an AFM image, and apply a smoothness constraint when drawing samples. The neural network-predicted ϵ_{NN} and h_{NN} are shown in Fig. 3(c)-(d). The predictions are averaged over 10 samples for each pixel. The predicted h_{NN} (Fig. 3(d)) shows a sharp edge on the left despite the smooth variation in the initial guess height, and ϵ_{NN} does not have a big variation, both agreeing with the physical expectations.

We condition the remaining parameters based on $(\epsilon_{\mathrm{NN}}, h_{\mathrm{NN}})$ from Fig. 3, using a modified probability weighted by $\exp(-\alpha D_j)$. Figure 5 displays sample predictions at a fixed optical wavelength of 759 nm and varying gate voltages. The bottom row depicts NN-predicted conductivity. At a large

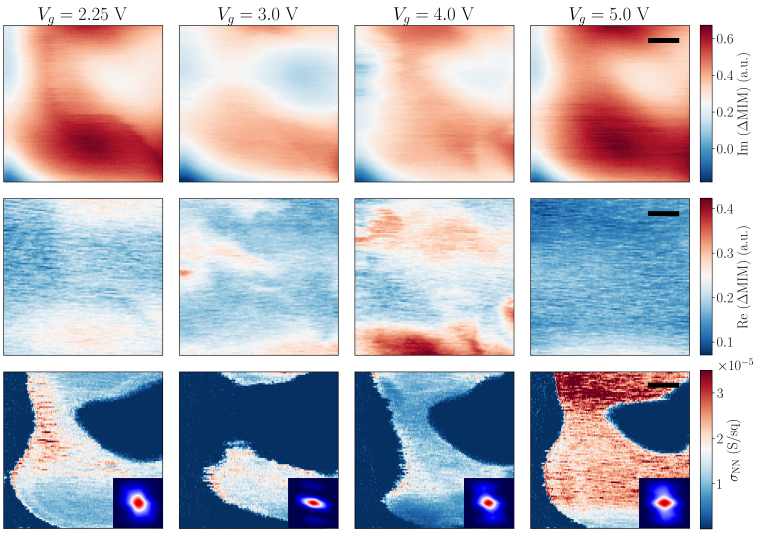

Figure 5: Imaginary (top) and real parts (middle) of the ER-MIM measurements at 4 different back gate voltages, V_g. From left to right: $V_g = 2.25$ V, 3.0 V, 4.0 V, and 5.0 V. The bottom row shows the neural network predicted conductivity. Bottom row insets: autocorrelation of the predicted conductivity. Scale bars are 200 nm.

voltage ($V_g = 5$ V), conductivity is large, which is consistent with semiconductor behavior under applied voltage. To quantitatively compare conductivity maps, we calculate the autocorrelation of σ_{NN} (insets in Fig. 5). At $V_g = 3$ V, A excitons are in resonance (not shown), leading to vertical autocorrelation suppression. We observed similar behavior for other gate voltages and resonant wavelengths, indicating that excitons enhance sample photoconductivity inhomogeneity.

4 Conclusions

We present an ML-based approach using ER-MIM measurements to predict material properties (sample topography, electrical and photoconductivity) and local environment (substrate permittivity). The probabilistic model incorporates physics knowledge, considering factors like a physical edge and sample inhomogeneity. Predicted conductivity and permittivity, along with exciton eigenenergy recombination [9], enable the prediction of nm-scale local electric field distribution, often inaccessible through traditional methods.

Broader Impact

In this work, we present a use case of an ML-based probabilistic model to accelerate the understanding of material properties. ML integration with scanning methods is only a recent development, with ER-MIM yet to benefit from it. The substantial data generated by these scans will likely lead to automated analysis in the near future. Our work represents a crucial stride toward automating the entire process, from data collection to property extraction, facilitating faster discovery of quantum phenomena and materials with reduced resource requirements.

Acknowledgments and Disclosure of Funding

Z.Z. and Z.J are supported by a Stanford Science Fellowship. Z.J. is also supported by the Urbanek-Chodorow fellowship.

References

[1] Mark E. Barber, Eric Yue Ma, and Zhi-Xun Shen. Microwave impedance microscopy and its application to quantum materials. *Nature Reviews Physics*, 4(1):61–74, 2022.

[2] Scott R. Johnston, Eric Yue Ma, and Zhi-Xun Shen. Optically coupled methods for microwave impedance microscopy. *Review of Scientific Instruments*, 89(4):043703, 04 2018.

[3] Oliver M Gordon and Philip J Moriarty. Machine learning at the (sub) atomic scale: next generation scanning probe microscopy. *Machine Learning: Science and Technology*, 1(2):023001, 2020.

[4] Nikolay Borodinov, Sabine Neumayer, Sergei V Kalinin, Olga S Ovchinnikova, Rama K Vasudevan, and Stephen Jesse. Deep neural networks for understanding noisy data applied to physical property extraction in scanning probe microscopy. *npj Computational Materials*, 5(1):25, 2019.

[5] Alexander Krull, Peter Hirsch, Carsten Rother, Augustin Schiffrin, and C Krull. Artificial-intelligence-driven scanning probe microscopy. *Communications Physics*, 3(1):54, 2020.

[6] Mohammad Rashidi and Robert A Wolkow. Autonomous scanning probe microscopy in situ tip conditioning through machine learning. *ACS nano*, 12(6):5185–5189, 2018.

[7] Xinzhong Chen, Suheng Xu, Sara Shabani, Yueqi Zhao, Matthew Fu, Andrew J Millis, Michael M Fogler, Abhay N Pasupathy, Mengkun Liu, and DN Basov. Machine learning for optical scanning probe nanoscopy. *Advanced Materials*, 35(34):2109171, 2023.

[8] Akash Laturia, Maarten L. Van de Put, and William G. Vandenberghe. Dielectric properties of hexagonal boron nitride and transition metal dichalcogenides: from monolayer to bulk. *npj 2D Materials and Applications*, 2(1):6, 2018.

[9] Deepankur Thureja, Atac Imamoglu, Tomasz Smoleński, Ivan Amelio, Alexander Popert, Thibault Chervy, Xiaobo Lu, Song Liu, Katayun Barmak, Kenji Watanabe, Takashi Taniguchi, David J. Norris, Martin Kroner, and Puneet A. Murthy. Electrically tunable quantum confinement of neutral excitons. *Nature*, 606(7913):298–304, 2022.