Research Article | Vol. 33, No. 7 / 7 Apr 2025 / Optics Express 14872
Optics EXPRESS

THz-TDS: extracting complex conductivity of two-dimensional materials via neural networks trained on synthetic and experimental data

Ben Beddoes,[1,*] Nicholas Klokkou,[1] Jon Gorecki,[2] Patrick R. Whelan,[3] Peter Bøggild,[3] Peter U. Jepsen,[4] and Vasilis Apostolopoulos[1,5,6]

[1]Physics and Astronomy, University of Southampton, Highfield SO17 1BJ, UK
[2]Department of Bioengineering, Imperial College London, London SW7 2AZ, UK
[3]Department of Physics, Technical University of Denmark, Kongens Lyngby DK-2800, Denmark
[4]Department of Electrical and Photonics Engineering, Technical University of Denmark, Kongens Lyngby DK-2800, Denmark
[5]University of Crete, Department of Physics, Heraklion GR-70013, Greece
[6]Foundation for Research and Technology-Hellas, Institute of Electronic Structure and Laser, Heraklion, Crete, Greece
*B.Beddoes@soton.ac.uk

Abstract: Terahertz time-domain spectroscopy (TDS) has proved immensely useful for probing 2D materials such as graphene. Unlike in the visible regime, the optical properties at terahertz frequencies are highly dependant on charge carrier mobility and scattering time. However, extracting the material properties from the terahertz waveform is a non-trivial process, which can be prone to producing erroneous results. Artificial neural networks have recently been demonstrated as useful tools to extract complex refractive index from terahertz time domain data. Here, we propose the use of artificial neural networks to interpret terahertz spectra of graphene monolayers to extract the charge carrier mobility and scattering time. We demonstrate improved performance on out-of-distribution data by using a combination of synthetically generated spectra and experimental data during training.

1. Introduction

Terahertz (THz) technologies and applications remain tantalisingly close to societal and commercial integration, with numerous studies highlighting the benefits of utilising this regime in sectors such as security, imaging, and communications [1–3]. The advantages of THz technologies are beginning to outweigh traditional drawbacks such as the hardware needed for generation and detection, and non-standardised analytical practices [4,5]. There is now a push towards applying these technologies to real-world applications. One industrial setting for which THz technology is particularly suited is understanding the complex conductivity of ultra-thin films such as 2D materials [6,7]. For example, graphene has seen significant progress in terms of fabrication [8–10] as well as device capability and performance [11]. However, there are currently few efficient, non-destructive methods for obtaining their electrical properties; existing methods often result in damage or contamination of the device during characterization [6]. THz time-domain spectroscopy (TDS), unlike many other spectroscopic tools, measures the amplitude of the electric field, preserving phase information and yielding high signal-to-noise ratios. The Fourier transform provides complete complex frequency-domain information and, through the division of a reference scan, allows complex parameters, including refractive index and conductivity as

#557580 https://doi.org/10.1364/OE.557580
Journal © 2025 Received 3 Feb 2025; revised 5 Mar 2025; accepted 16 Mar 2025; published 25 Mar 2025

Optics EXPRESS

well as other key performance factors such as carrier density and mobility, to be derived [12–14]. THz-TDS has the added benefit of being a fast, non-contact, non-destructive characterization technique that can be applied to arbitrarily large sample areas while preserving device integrity [15].

Yet, obtaining material parameters requires multiple analysis stages including: phase unwrapping, jitter correction, and transfer function convergence, with each step introducing potential errors [16,17]. Much work has been carried out to improve the robustness of THz-TDS extraction, whether through the standardization of practices [5] or addressing individual extraction issues [18]. For example, introducing a timing jitter correction solves an issue where even small time shifts can cause significant variations in the electrical properties [19]. However, standardization has yet to reduce the number of pre-processing steps, and incorporating error-reducing methods into traditional fitting algorithms results in a more computationally demanding extraction process [6], which is less suitable for industrial-scale testing.

Artificial neural networks (ANNs) can assist in several key ways. Firstly, they have already been shown to equal or even surpass traditional fitting techniques in terms of accuracy for parameter extraction, see [20] and references therein. ANNs reduce user intervention during THz-TDS extraction by cutting the number of variables and decisions during processing, with recent studies showing that it is even possible to extract parameters from the time-domain pulses themselves without any need for Fourier analysis [21]. Additionally, ANNs offer the advantage of fast implementation and convergence when compared to traditional methods, making them prime candidates for industrial-scale testing and real-world application [22,23]. The performance of ANNs is strongly tied to the availability of data. For supervised learning with real-world data this requirement is difficult to satisfy in the THz regime where large experimental datasets can be difficult to obtain. Our work, which began in 2020, first showed ANNs trained with purely synthetic data could, under certain circumstances, be as effective and quicker than the traditional Newton-Raphson fit at obtaining the complex refractive index from slab samples [22]. Using synthetic data allowed us to bypass the traditional drawback of ANNs, as correctly labelled large synthetic datasets are comparatively easy to create. Since then, we have shown ANNs trained with synthetically generated data can be robust against noise and phase offsets [24]. Synthetically trained ANNs are still fundamentally limited by the applicability of the model used to train them, adding artificial noise can only go so far to capture the real-world system dynamics of experimentally obtained data and bridging this gap is seen as a key step in the progress of ANNs [25].

In this work, we investigate the performance of ANNs trained to extract the complex conductivity of graphene monolayers from experimental THz-TDS measurements. The networks take inputs of frequency-domain data and output values for the DC conductivity, σ_{DC}, and carrier scattering time, τ, from which other properties such as carrier density and drift velocity can easily be inferred. We present three training regimes: the first trained exclusively with synthetic data; the second trained exclusively with experimental data; and the final network trained with a combination of both. Our results demonstrate how training with both experimentally obtained data and synthetically generated data can enhance the accuracy, generality, and robustness of the network to real-world dynamics as well as reduce the required amount of real-world data to train an effective network. We aim to build upon this work by obtaining experimental data from a range of sources and simulating data using other models, such as the Drude-Smith model, to increase the overall applicability and generality of our networks by capturing a broader range of behaviours, extending to other 2D materials.

2. THz-TDS conductivity extraction of 2D materials

THz-TDS has proven to be an effective tool for investigating the complex conductivity of 2D materials including semiconductors, nanomaterials and graphene [26–28]. The complex

Optics EXPRESS

conductivity can be obtained using THz-TDS by measuring the attenuation and phase delay of the THz pulse through the sample in either reflection or transmission mode. Most TDS systems utilise a femtosecond laser, which is split between two paths: one directed at a THz emitter and the other at a THz detector. A delay line is incorporated into one of the arms to allow the capture of the full time-domain pulse. In this work, a commercial fiber-coupled THz-TDS system (from Picometrix, Inc.) was used on a high-resistivity silicon wafer with a graphene monolayer deposited on top. The sample is grown by conventional chemical vapor deposition process [29], involving decomposition of methane onto a copper thin film to form polycrystalline monolayer graphene. This is followed by deposition of a polymethyl methacrylate (PMMA) carrier film, which is then moved onto a high-resistive silicon wafer with a silicon dioxide top layer, to enable THz-TDS pulses to penetrate the wafer without attenuation. Finally, the polymer is removed by organic solvents, leaving the graphene on the Si/SiO_2 wafers. The wafer is mounted on an xy stage, see Fig. 1(a), and was raster-scanned within the focal plane of the THz beam to create spatial maps, with a resolution of 300 μm at 1 THz [19]. In these scans, areas without graphene serve as a reference during parameter extraction, as shown in the example scan in Fig. 1(b).

The advantage of such a system over traditional techniques, such as optical or electron beam lithography, which involve contact with polymer resists, etching, and metal deposition, or measurements with fixed or movable probes, which require direct physical contact [6], lies in its ability to operate without contact. THz TDS does not alter the sample, allowing it to be used for its intended purpose following characterization. Figure 1(c) and (d) present typical values for the DC conductivity, σ_{DC}, and the carrier scattering lifetime, τ, of thin graphene layers extracted from THz TDS data using an iterative root finding method. Such extractions typically involve the Newton-Raphson or Nelder-Mead method, which is regarded as the gold standard for THz-TDS analysis [30–33]. Here, we have used MATLAB's nonlinear least-squares solver.

Extracting the complex conductivity of thin films from THz-TDS data involves mapping a theoretical transfer function to an experimentally acquired transfer function [13,34,35]. Our focus is on transmission measurements, with samples that satisfy the thin-film limit ($d << \frac{\lambda}{2\pi n}$), and where the conducting layer dominates the optical properties of the sample due to the strong interaction of free carriers with the incident THz field [6]. The experimental transfer function is obtained by measuring the electric field of the THz time-domain pulses transmitting through the aforementioned graphene coated (sample) and bare silicon (reference) regions. After several pre-processing steps outlined in Fig. 2, including time-domain windowing to remove echos, performing a Fast Fourier transform (FFT), and phase reconstruction, the complex, frequency dependent experimental transfer function is calculated,

$$\tilde{T}_{\mathrm{Exp}}(\omega) = \frac{\tilde{E}_{\mathrm{Sam}}(\omega)}{\tilde{E}_{\mathrm{Ref}}(\omega)}, \tag{1}$$

where $\tilde{E}_{\mathrm{Sam}}(\omega)$ and $\tilde{E}_{\mathrm{Ref}}(\omega)$ correspond to the frequency-domain electric field of the sample and reference scans, respectively. By referencing in this way the spectra amplitude becomes normalised. A theoretical transfer function is then mapped to $\tilde{T}_{\mathrm{Exp}}$, derived from the corresponding theoretical descriptions of reference and sample electric fields. For the conductive layer, we introduce the transmission coefficient,

$$\tilde{t}_{\mathrm{air,film,sub}}(\omega) = \frac{2}{1 + n_{\mathrm{sub}} + Z_0\tilde{\sigma}(\omega)}, \tag{2}$$

which represents the propagation of the THz pulse from air, through the conductive film and into the substrate. Consequently, $\tilde{t}_{\mathrm{air,film,sub}}$ depends on the refractive index of the substrate (n_{sub}), the impedance of free space ($Z_0 = \sqrt{\frac{\mu_0}{\epsilon_0}} = 377\ \Omega$), and the complex conductivity ($\tilde{\sigma}(\omega)$). The combination of the transmission coefficient, given by Eq. (2), and the Fresnel coefficients (not

Optics EXPRESS

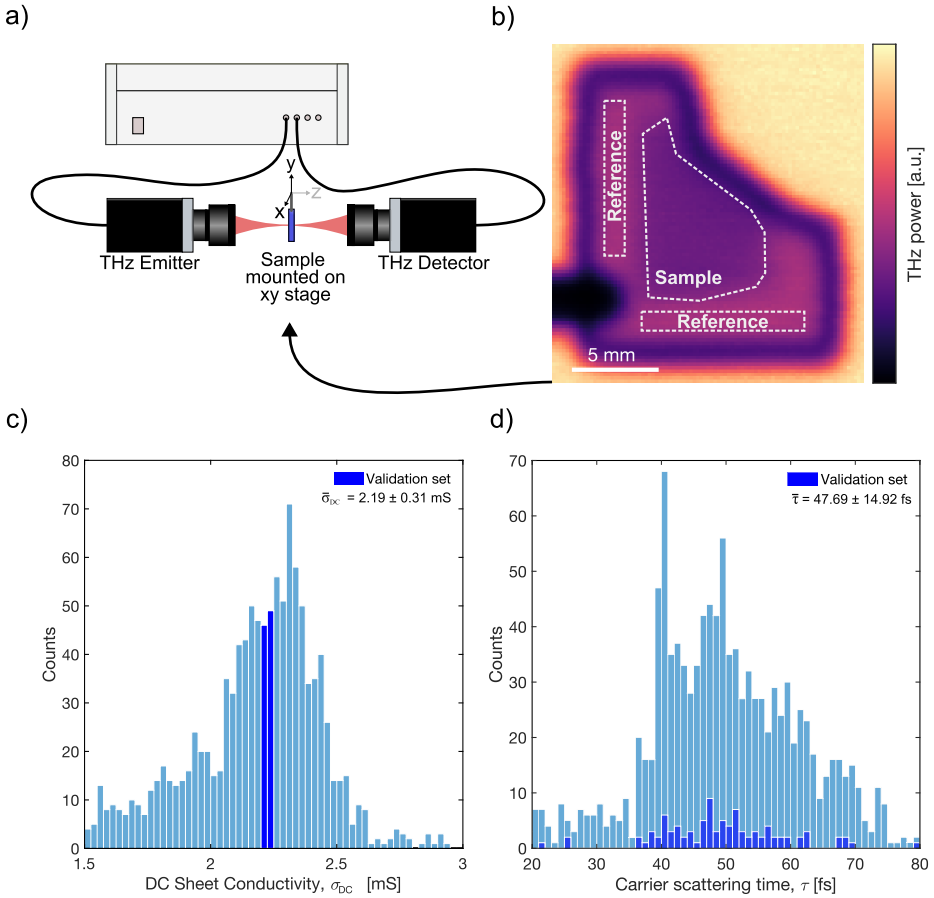

Fig. 1. Experimental data used for this work, acquired from an imaging system based on a commercial spectrometer (a), which produces the power map (b). Each pixel in the map has a full time-domain pulse, and areas are selected for reference (bare substrate) and sample (graphene coated), identified by the relative attenuation. Performing the parameter extraction and Drude model fitting for each pixel results in a range of values for DC conductivity (c) and scattering time (d). The values highlighted in dark blue correspond to the spectra values of DC conductivity and scattering time that have been removed for the out-of-distribution (OOD) performance characterization.

Optics EXPRESS

shown) for both the sample and reference, yield the theoretical transfer function for the direct pass of the THz pulse (Tinkham formula [36]),

$$\tilde{T}^{(0)}_{\text{Theo}}(\omega) = \frac{1 + n_{\text{sub}}}{1 + n_{\text{sub}} + Z_0\tilde{\sigma}(\omega)}. \tag{3}$$

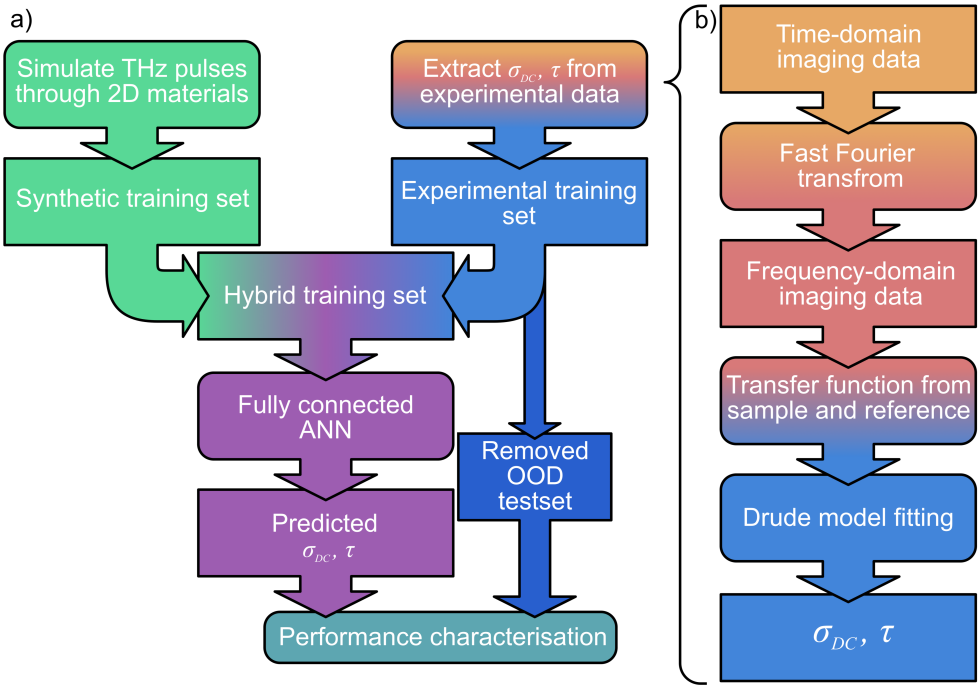

Fig. 2. Flowchart describing (a) the machine learning workflow for the hybrid training approach and (b) the parameter extraction steps for THz-TDS measurements. Synthetic data is generated by simulating THz pulses through 2D materials with a range of σ_{DC} and τ values. An experimental dataset is produced by extracting parameters from experimental data using a fitting method. An ANN is then trained on a mixture of both datasets resulting in a *hybrid trained network*. Taking only the left or right path for dataset generation results in a *synthetic network* and an *experimental network* respectively. A subset of experimental data is withheld from the training step and used for the out-of-distribution (OOD) performance characterization. The parameter extraction process (b) has various steps that can result in errors such as time-domain windowing and phase retrieval.

Several models can be used to fit $\tilde{\sigma}$; for this report, we focus on the most commonly used, the Drude model, given by

$$\tilde{\sigma}(\omega) = \frac{\sigma_{\text{DC}}}{1 - i\omega\tau}, \tag{4}$$

where σ_{DC} represents the sheet DC conductivity, and τ is the carrier scattering time, from which other parameters, including the carrier density and drift velocity, can be inferred. To find a solution for σ_{DC} and τ, an iterative root-finding or minimization algorithm is employed to directly fit to the Drude model. Despite its widespread use, this method can converge to local or non-global solutions. Another issue is the computation time which can become a significant issue for industrial scalability. Other models, including the Drude-Smith formulations, exist for samples that do not follow Drude-like behaviour [15]. In these situations, it is up to the user to select a more appropriate model for their target material. It is for these reasons we look to artificial neural networks (ANNs) as a possible solution.

Optics EXPRESS

3. Generating synthetic training sets

There has been a recent upsurge in the use of ANNs to analyse THz-TDS data in an attempt to reduce the amount of processing steps and user input during parameter extractions. This includes the refractive index measurements of slab samples from synthetic data [22], leaf wetness for predicting plant disease [37], and the conductivity of nanowires using experimental data [38]. In each case, significant gains have been made in both accuracy and speed of the extraction process, yet the training data has come from a singular source, i.e purely experimental or synthetic data. In this section we demonstrate the capability of synthetically trained data to perform comparably to traditional iterative root finding methods, before further improvements are gained in the next section by including experimentally obtained data into the training sets. All work in this report has been carried out on a Dell Optiplex 7090 with an Intel i7-10700 processor, 24 GB of RAM and a NVIDIA GeForce GTX 1660 GPU.

To solve conductivity problems with machine learning algorithms, we aim to address the theoretical transfer function given in Eq. (3) by generating synthetic data. We follow a similar method established in [22], where a characteristic terahertz time-domain pulse is modelled and propagated through a reference material, in this case, silicon, modelled using Fresnel transmission coefficients, $t_{x,y}$. The synthetically generated electric field of the transmitted reference signal is given by

$$\tilde{E}^{(0)}_{\mathrm{SynRef}}(\omega) = \tilde{E}_0(\omega) \cdot t_{\mathrm{air,sub}}(\omega) \cdot P(\omega) \cdot t_{\mathrm{sub,air}}(\omega). \tag{5}$$

Similarly, a second time-domain pulse is modelled and propagated through a sample using the Fresnel coefficients and transmission relation from Eq. (2). To model the sample propagation we generate random values for σ_{DC} and τ within a specified range: σ_{DC} ranges from 0.3 to 10 mS and τ ranges from 1 to 100 fs. These ranges encompass typical values for thin graphene samples (see Fig. 1(c,d)). These values are fed into the Drude model, see Eq. (4), to give the complex conductivity $\tilde{\sigma}(\omega)$ used in $\tilde{t}_{\mathrm{air,film,sub}}$. The synthetically generated electric field of the transmitted sample signal is given by

$$\tilde{E}^{(0)}_{\mathrm{SynSam}}(\omega) = \tilde{E}_0(\omega) \cdot \tilde{t}_{\mathrm{air,film,sub}}(\omega) \cdot P(\omega) \cdot t_{\mathrm{sub,air}}(\omega). \tag{6}$$

P here is the propagation through the layer and depends upon the substrate refractive index, n_{sub} and thickness, d. Silicon is an ideal substrate for the THz-TDS (0.1 - 3 THz) as it has low absorption and a non-dispersive refractive index [39].

$$P(\omega) = e^{-idn_{\mathrm{sub}}\frac{\omega}{c}} \tag{7}$$

The division of $\tilde{E}^{(0)}_{\mathrm{SynSam}}$ with $\tilde{E}^{(0)}_{\mathrm{SynRef}}$ will yield the synthetically generated transfer function, effectively reversing the previously described parameter extraction. This also keeps the normalization consistent between experimental and synthetically generated datasets. The synthetically generated spectra accompanied by the values of σ_{DC} and τ serve as our training set and target values for our ANN.

3.1. Training with large synthetic datasets

For our ANNs each frequency point has two neurons corresponding to values of amplitude and phase. The target output comprises values for σ_{DC} and τ. To reduce computation and training overhead, the frequency data is down-sampled into 32 points, providing the network with 64 input values. We artificially added 60 dB of random noise (in the frequency domain) to the training sets which consisted of 200,000 synthetically generated spectra. The network architecture had nine hidden layers, with four layers of 33 and five layers of 11 fully connected neurons. We used a supervised training method from MATLABs deep learning toolbox to predict a known output array, using the Levenberg-Marquardt algorithm. The activation function is a sigmoid in

Optics EXPRESS

the hidden layers and a linear function in the output layer. A frequency range of 0.3 to 2 THz was selected corresponding to the highest SNR for the spectrometer. It is worth noting that the material parameter and frequency ranges can be readily extended to include other conductive samples and higher bandwidth spectrometers.

Figure 3 presents a comparison between a conventional Drude fit and our synthetically trained network. We used MATLAB's nonlinear least squares solver to fit the Drude model to the complex conductivity data. This method requires truncating the frequency axis to a range of 0.3 to 1.5 THz in order to obtain reasonable values, ensuring that the data has a sufficiently high signal-to-noise ratio (SNR). THz-TDS spectra have the lowest SNR at high and low frequencies with the best SNR usually around 1 THz, although this strongly depends on the spectrometer used. In regions of low SNR iterative methods have a tendency not to converge to a solution [22]. The need to restrict the frequency range highlights an advantage of the network approach, as it does not require data truncation for stable parameter extraction even in regions of lower SNR. The objective function for the solver is complex and represents the difference between the complex conductivity derived from the experimentally obtained transfer function (Eq. (3), rearranged for σ_{DC}), and that predicted by the Drude model in Eq. (4), which is only modified to ensure that both σ_{DC} and τ are real numbers. For the experimentally obtained data, the iterative Drude method returned values of $\sigma_{DC} = 2.56$ mS and $\tau = 61.71$ fs. In comparison, our network produced values of $\sigma_{DC} = 2.49$ mS and $\tau = 56.3$ fs. This demonstrates that our synthetic network performs comparably to traditional fitting methods and performs better when taken over the full frequency range. This network was able to predict the values of σ_{DC} and τ to an overall error of 6.3% when compared to the traditional Drude extraction for 100 unseen experimental spectra. The iterative method also took longer to fit the spectra, 0.042 ± 0.005 s compared to the network execution time of 0.022 ± 0.006 s.

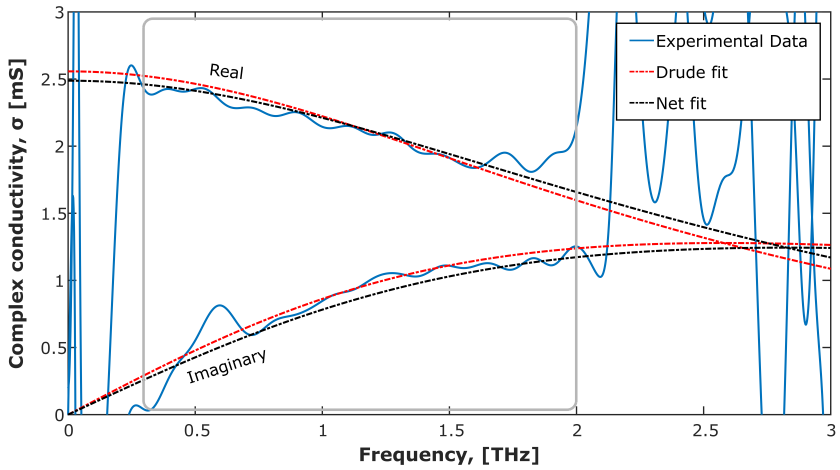

Fig. 3. Real and imaginary parts of the complex conductivity for experimental data (blue line), Drude fit (red line) and network fit (black line). The Drude fit was calculated using the data range 0.3–1.5 THz. The net fit was calculated using the data range 0.3–2 THz. The grey box represents the range of usable data to which a Drude fit can be made. Data outside the grey box is excluded from all analysis.

4. Network enhancement using hybrid training sets

To further improve the accuracy and generalisability of our networks, in this section we present three new networks which underpin the benefits of mixing experimentally obtained data with

Optics EXPRESS

synthetic data in the network training sets. The first network has been trained using exclusively synthetic data, the second exclusively experimental data, and the third a 50:50 mix of experimental and synthetic data which will now be referred to as the hybrid trained network. To give a fair comparison and focus on the effect of combining synthetic and experimental data, the networks presented in this section have all been trained using the same algorithm, the same network architecture of [32 x 6 x 6] fully connected hidden layers, and using a frequency range of 0.3 - 1.5 THz. These changes were made to compensate for the limited availability of experimental data. We have reduced the amount of synthetic data used to train the network to 1600 datasets, which resulted in the change in network architecture to that described previously. The frequency range was changed so that the Drude fitting method used was able to produce accurate target values for σ_{DC} and τ which the experimental network will use to train. The exclusively experimental network has been trained with 800 datasets while the hybrid trained network has 800 datasets of each for a total of 1600. Each synthetic training set has been augmented with artificial noise as described previously [24]. Otherwise the method remains the same as described in section 3.1.

Prior to training a range of experimental data corresponding to values of σ_{DC} in the range 2.2 - 2.25 mS was removed from all experimental training sets. This ensures that each network has not seen experimental data within this range and provides us with a fair test set to compare network performance. These spectra correspond to the values highlighted in Fig. 1(c) and (d). This method is referred to as out-of-distribution (OOD) data generalization and is a method of testing that ensures that each network is evaluated on data that lies outside the parameters of the training sets [40,41]. This means the features, patterns, or statistical properties of the OOD data differ from the training data allowing us to assess the robustness and generalisability of each network to unseen scenarios that have not been encountered during training. The synthetic training sets will have seen σ_{DC} values within this range; however, lack experience of real-world dynamics. This ensures that we can still assess the generalisability and performance of the network using this method. The same method was used to remove a range of τ values with similar results and are therefore not presented here.

Since the ANNs used in this report are trained with different combinations of synthetic and experimental data, we found that MATLAB's built-in network validation (which uses mean squared error (MSE)) can be misleading when assessing network accuracy. This method involves setting aside 15% of the original dataset, ensuring that the network never encounters this data during training. After training, the network's performance is then evaluated using this unseen data. Using this MSE metric, the results for each network were as follows: synthetic network: 1.79×10^{-5}; experimental network: 1.28×10^{-4}; hybrid trained network: 1.24×10^{-4}. Based solely on these performance metrics, one might expect the purely synthetic network to significantly outperform both the experimental and hybrid trained networks. However, this validation method assesses only the data type used in training each network. In other words, the synthetic network's validation metric reflects its ability to extract information from unseen synthetic data, which does not provide a fair or meaningful comparison across all networks. To ensure a more robust and objective evaluation, we have validated each network's performance on an unseen experimental OOD dataset.

An additional 1000 synthetically generated spectra was created to provide a second test set for our networks. These synthetically generated spectra covered the full data range. We have plotted the network values against the target values obtained via traditional fitting methods for both the experimental OOD range (blue circles) and external synthetic test set (red circles). For a successful network we expect data points close to the black line which represents a perfect match between the network and target values. All graphs for σ_{DC} and τ have been plotted with the same axes for comparison.

Optics EXPRESS

4.1. Network training set: synthetic

The use of synthetic data to train networks to model real-world data is both highly convenient and effective. It enables the rapid and straightforward generation of large datasets, which would otherwise require considerable time and effort to produce experimentally [42]. Synthetically trained networks are by default more general than their experimental counterparts as a synthetically trained network should be able to extract parameters regardless of the spectrometer used. This issue is especially applicable to THz study, where variations in fitting algorithms and experimental setups has led to non-standardised practices and variations in extracted material parameters for even simple samples between different spectrometers [16,22]. Synthetic data can also be generated for arbitrary values of σ_{DC} and τ, allowing us to supplement our training sets, whether because of a lack of experimental data or to extract from new samples.

The results for the synthetic network are presented in Fig. 4(a) (I-III). This network uses a training set comprised of 1600 synthetic spectra covering the full data range as shown in Fig. 4(a) (I). The network has then been tested using the external synthetic test set for which it was able to predict values of σ_{DC} and τ to an error of 1.4%, see red circles in Fig. 4(a) (II, III). We would expect this network to perform well against the external synthetic test set as both sets have been generated with the same method. For the OOD experimental test set the network is able to obtain values for σ_{DC} and τ with a 10.01% error of the Drude fit (blue circles). It is unsurprising that the network performs worse than the large dataset synthetic network described in section 3.1, as the amount of data available generally tracks with network performance. There is also a tendency to underestimate σ_{DC}, which is most likely caused by underfitting due to a lack of available data [43]. Despite these issues the network provides us with a good benchmark of performance with which to compare the subsequent networks. Despite the benefits of using synthetic networks, fundamentally the accuracy is still limited by the effectiveness of the model used for training, in this case the Drude model. As mentioned, our networks include artificial noise and errors to try to capture real-world dynamics within the training sets, like jitter or phase offsetting. However, capturing all dynamics of a system is difficult artificially, which is why experimentally acquired data is normally the first consideration. Another possible solution is to train networks using different models for conductivity, i.e the Drude-Smith model. This would allow us to include new dynamics and behaviours within our training sets, improving the applicability.

4.2. Network training set: experimental

Training on real-world data is advantageous for the obvious reason that all system dynamics are incorporated into the training sets. Given an arbitrarily large dataset, it would be possible to create an arbitrarily accurate network. However, obtaining real-world datasets is time-consuming and not always feasible. Furthermore, the THz-TDS-specific issue of variability in experimental results obtained from different spectrometers necessitates data from a range of sources to prevent the network from overfitting or learning specific features or faults of a particular setup or parameter extraction method. This further increases the complexity and time required to acquire sufficiently large datasets. Although the time required to characterise wafer samples is continuously improving, for example, over the past decade the extraction of 12 inch wafers using THz-TDS has decreased from hours to minutes, gathering datasets of suitable size would still be a significant undertaking [15].

The results presented in Fig. 4(b) (I-III) are for the network trained using exclusively experimental data. The data used came from the same source. The grey highlighted region in Fig. 4(b) (I) represents the range of σ_{DC} data removed for OOD testing. The network will never see data within this region during training. This network fails to predict values of σ_{DC} and τ for the external synthetic test set, hence few red circles exist near to the perfect match line (black line) in Fig. 4(b) (II-III). However, there is a noticeable improvement in accuracy over the purely synthetic network for the OOD experimental test set, despite having only half the number of

Optics EXPRESS

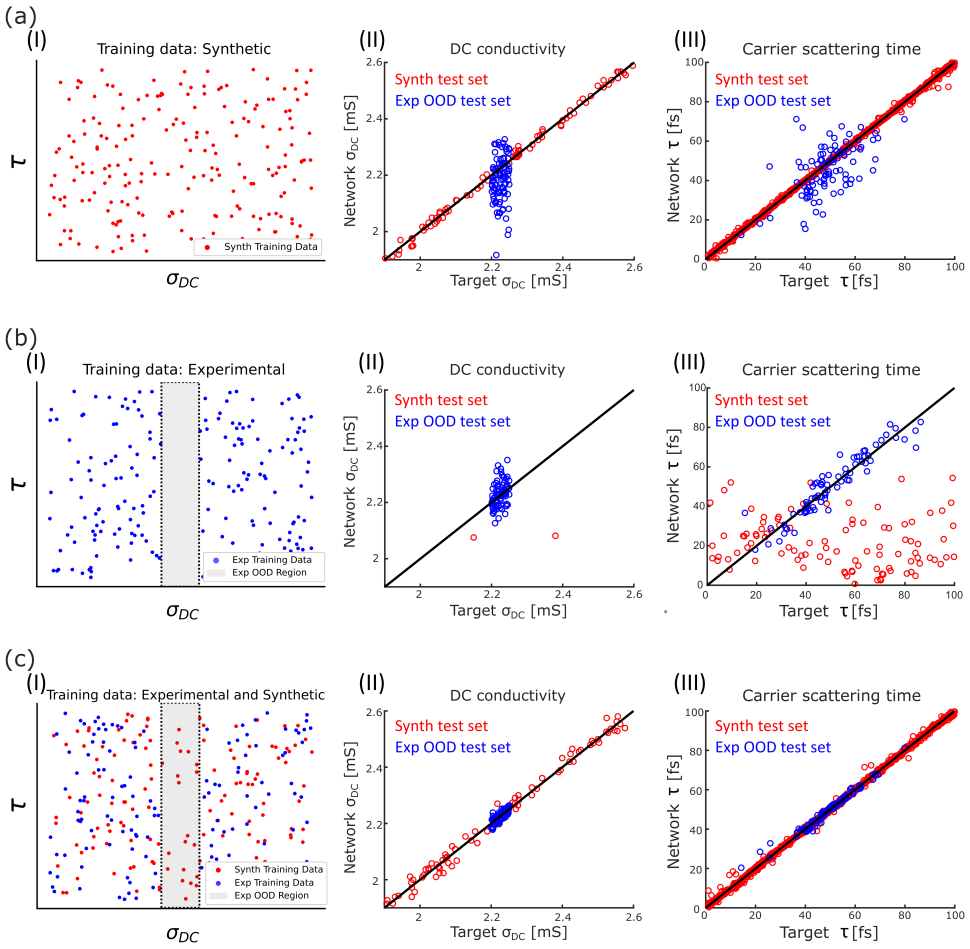

Fig. 4. Results for the three networks: (a) Trained with 1600 synthetic training sets, (b) trained with 800 experimental training sets, and (c) trained with a 50:50 mix of 800 synthetic and 800 experimental training sets. (I) Visualization of the data used to train each network. Synthetic training data is indicated by red dots, and experimental data by blue dots. The grey region represents the out-of-distribution (OOD) region, where experimental data has been removed to create an external test set unseen by any network during training. (II) Network σ_{DC} predictions compared to synthetic target values and Drude fit values for the external synthetic test set (red circles) and the experimental OOD test set (blue circles). The black line represents a perfect match between target and network predictions. (III) Network τ predictions compared to synthetic target values and Drude fit values for the external synthetic test set (red circles) and the experimental OOD test set (blue circles). The black line represents a perfect match between target and network predictions. All three networks can predict σ_{DC} and τ from the OOD test set; however, the best performance is achieved with a mixed training set, as shown by data points close to the perfect match line (black line). For the synthetic test set, similar performance is observed with synthetic or mixed networks, while the exclusively experimental network fails to return reasonable values.

Optics EXPRESS

datasets. The network is able to predict values of σ_{DC} and τ to an error of 4.77% of the quoted value for the OOD experimental test set. We would expect some improvement over the synthetic network as it has seen data from this exact spectrometer above and below the removed values of σ_{DC}. The data will also contain more of the real-world dynamics compared to the synthetic training sets. Further improvements to the network would require larger datasets. Augmentation of the experimental data in manner similar to that used with the synthetic network e.g: artificially including noise, can be used to increase the sample size, however this does not solve the problem of overfitting to one particular system.

4.3. Network training set: hybrid

For these reasons, we look to include a combination of experimental and synthetic spectra in our training sets. The hybrid trained network presented in Fig. 4(c) (I-III) uses 800 synthetic spectra, which cover the full range of σ_{DC} and τ values and 800 experimentally obtained training sets which, as previously, have had values of σ_{DC} removed. This is represented in Fig. 4(c) (I) with the grey highlighted box showing the OOD experimental region, however unlike for the purely experimental network, synthetic training spectra do lie within this region. This network is able to predict values of σ_{DC} and τ for the external synthetic test set to an error of 2.11% compared to the target values. This is comparable to the synthetic networks trained previously. For the OOD experimental test set the network was able to predict values with an error of 1.70% compared to the Drude fit. This is a significant improvement over the previous two methods. Furthermore, this is achieved using the same amount of experimental training sets.

To provide an estimate of the amount of experimental data required to achieve an extraction accuracy of 1.7% (the best achieved by our hybrid trained network), we trained purely experimentally networks with progressively fewer data sets and recorded the corresponding accuracy. By fitting this data with a power-law function, we estimate that at least 10,000 experimental datasets would be needed to match the accuracy of our hybrid trained network. This is approaching an order of magnitude higher than the number of datasets required for the hybrid training approach. We also examined how altering the ratios of experimental and synthetic data influenced extraction. Even small additions ($\approx$ 20 %) of synthetic data to experimental training sets significantly enhanced performance.

We show that the inclusion of synthetic data within the network training set can not only supplement for less widely available experimental data, but also improve network performance against real-world OOD data. For the case presented here, we are able to reduce the error from given target values by a third when compared to a network trained exclusively with experimental data. Through no increase in experimental data we are able to better capture real-world dynamics of 2D systems whilst also maintaining the generalisability of synthetically trained networks. We anticipate enhanced performance by investigating further the ratio of experimental to synthetic data, optimising network methods, developing strategies to incorporate data from a range of sources, and increasing dataset size. Further improvements will be achieved in our synthetic training sets through multiple conductivity models and the inclusion of physics-informed loss functions [44].

5. Conclusion

In this work, we have presented a method for generating large-scale datasets for conductive 2D materials based on the Tinkham relations. We demonstrated that an artificial neural network, trained using a large dataset of exclusively synthetic spectra, performed comparably to traditional fitting methods when extracting graphene parameters from real-world THz-TDS data, and even outperformed them when used over the full spectrometer bandwidth. We then investigated how the network's performance could be further improved by using a combination of synthetic and experimental training sets, demonstrating marked improvements in accuracy over networks

Optics EXPRESS

trained on a singular source. This method allows us to reduce the demand for experimental data by supplementing training sets with synthetic data. Additionally, network performance is enhanced by combining the generality provided by synthetic training sets with the robustness to the dynamics of real-world systems achieved through training on experimental data, all while maintaining fast extraction times. The focus of our work is to expand our training sets to incorporate different conductive systems by increasing the size and diversity of our experimental datasets, including data from other sources and materials. Additionally, we plan to integrate other models, such as the Drude-Smith model, into our synthetic training sets to further improve the generality and applicability of our networks, preparing them for use in industrial-scale testing. This integration could either complement traditional fitting algorithms or serve as a complete replacement for them.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are available from the University of Southampton repository at [45].

References

1. Y. Cheng, L. Qiao, D. Zhu, *et al.*, "Passive polarimetric imaging of millimeter and terahertz waves for personnel security screening," Opt. Lett. **46**(6), 1233–1236 (2021).
2. Y. Huang, Y. Shen, and J. Wang, "From terahertz imaging to terahertz wireless communications," Engineering **22**, 106–124 (2023).
3. S. Shi, S. Yuan, J. Zhou, *et al.*, "Terahertz technology and its applications in head and neck diseases," iScience **26**(7), 107060 (2023).
4. A. Leitenstorfer, A. S. Moskalenko, T. Kampfrath, *et al.*, "The 2023 terahertz science and technology roadmap," J. Phys. D: Appl. Phys. **56**(22), 223001 (2023).
5. J. Lee, C. K. Leung, M. Ma, *et al.*, "The dotthz project: A standard data format for terahertz time-domain data," J. Infrared, Millimeter, Terahertz Waves **44**(11-12), 795–813 (2023).
6. P. Bøggild, D. M. Mackenzie, P. R. Whelan, *et al.*, "Mapping the electrical properties of large-area graphene," 2D Mater. **4**(4), 042003 (2017).
7. M. Nagai, S. Watanabe, R. Imamura, *et al.*, "Characterization of ultrathin conductive films using a simplified approach for terahertz time-domain spectroscopic ellipsometry," J. Infrared, Millimeter, Terahertz Waves **45**(11-12), 949–966 (2024).
8. V. B. Mbayachi, E. Ndayiragije, T. Sammani, *et al.*, "Graphene synthesis, characterization and its applications: A review," Res. Chem. **3**, 100163 (2021).
9. F. Bonaccorso, A. Lombardo, T. Hasan, *et al.*, "Production and processing of graphene and 2d crystals," Mater. Today **15**(12), 564–589 (2012).
10. D. Kireev, D. Sarik, T. Wu, *et al.*, "High throughput transfer technique: Save your graphene," Carbon **107**, 319–324 (2016).
11. A. C. Samaha, J. Doumani, T. E. Kritzell, *et al.*, "Graphene terahertz devices for sensing and communication," Small p. 2401151 (2024).
12. J. Neu and C. A. Schmuttenmaer, "Tutorial: An introduction to terahertz time domain spectroscopy (thz-tds)," J. Appl. Phys. **124**(23), 231101 (2018).
13. J. Lloyd-Hughes and T. I. Jeon, "A review of the terahertz conductivity of bulk and nano-materials," J. Infrared, Millimeter, Terahertz Waves **33**(9), 871–925 (2012).
14. T. D. Dorney, R. G. Baraniuk, and D. M. Mittleman, "Material parameter estimation with terahertz time-domain spectroscopy," J. Opt. Soc. Am. A **18**(7), 1562–1571 (2001).
15. P. R. Whelan, B. Zhou, O. Bezencenet, *et al.*, "Case studies of electrical characterisation of graphene by terahertz time-domain spectroscopy," 2D Mater. **8**, 022003 (2021).
16. M. Naftaly, "An international intercomparison of thz time-domain spectrometers," *41st International Conference on Infrared, Millimeter, and Terahertz Waves* pp. 1–2 (2016).
17. P. U. Jepsen, "Phase retrieval in terahertz time-domain measurements: a "how to" tutorial," J. Infrared, Millimeter, Terahertz Waves **40**(4), 395–411 (2019).
18. P. R. Whelan, Q. Shen, D. Luo, *et al.*, "Reference-free thz-tds conductivity analysis of thin conducting films," Opt. Express **28**(20), 28819 (2020).
19. P. R. Whelan, K. Iwaszczuk, R. Wang, *et al.*, "Robust mapping of electrical properties of graphene from terahertz time-domain spectroscopy with timing jitter correction," Opt. Express **25**(3), 2725–2732 (2017).
20. M. Gezimati and G. Singh, "Terahertz data extraction and analysis based on deep learning techniques for emerging applications," IEEE Access **12**, 21174–21198 (2024).
21. M. Z. Güngördü, P. Kung, and S. M. Kim, "Non-destructive evaluation and fast conductivity calculation of various nanowire-based thin films with artificial neural network aided thz time-domain spectroscopy," Opt. Express **31**(6), 10657 (2023).

Optics EXPRESS

22. N. Klokkou, J. Gorecki, J. S. Wilkinson, *et al.*, "Artificial neural networks for material parameter extraction in terahertz time-domain spectroscopy," Opt. Express **30**(9), 15583 (2022).
23. L. Fan, Y. Zeng, Q. Yang, *et al.*, "Fast and high-quality 3-d terahertz super-resolution imaging using lightweight sr-cnn," Remote Sens. **13**(19), 3800 (2021).
24. N. Klokkou, J. Gorecki, B. Beddoes, *et al.*, "Deep neural network ensembles for thz-tds refractive index extraction exhibiting resilience to experimental and analytical errors," Opt. Express **31**(26), 44575 (2023).
25. T. Alkhalifah, H. Wang, and O. Ovcharenko, "Mlreal: Bridging the gap between training on synthetic data and real data applications in machine learning," Artif. Intell. Geosci. **3**, 101–114 (2022).
26. A. M. Ulatowski, L. M. Herz, and M. B. Johnston, "Terahertz conductivity analysis for highly doped thin-film semiconductors," J. Infrared, Millimeter, Terahertz Waves **41**(12), 1431–1449 (2020).
27. H. J. Joyce, J. L. Boland, C. L. Davies, *et al.*, "A review of the electrical properties of semiconductor nanowires: Insights gained from terahertz conductivity spectroscopy," Semicond. Sci. Technol. **31**(10), 103003 (2016).
28. P. R. Whelan, D. Huang, D. Mackenzie, *et al.*, "Conductivity mapping of graphene on polymeric films by terahertz time-domain spectroscopy," Opt. Express **26**(14), 17748 (2018).
29. J. D. Buron, D. H. Petersen, P. Bøggild, *et al.*, "Graphene conductance uniformity mapping," Nano Lett. **12**(10), 5074–5081 (2012).
30. I. Pupeza, R. Wilk, and M. Koch, "Highly accurate optical material parameter determination with thz time-domain spectroscopy," Opt. Express **15**(7), 4335–4350 (2007).
31. J. C. Lagarias, J. A. Reeds, M. H. Wright, *et al.*, "Convergence properties of the nelder–mead simplex method in low dimensions," SIAM J. Optim. **9**(1), 112–147 (1998).
32. L. Duvillaret, F. Garet, and J.-L. Coutaz, "Highly precise determination of optical constants and sample thickness in terahertz time-domain spectroscopy," Appl. Opt. **38**(2), 409–415 (1999).
33. L. Duvillaret, F. Garet, and J.-L. Coutaz, "A reliable method for extraction of material parameters in terahertz time-domain spectroscopy," IEEE J. Sel. Top. Quantum Electron. **2**(3), 739–746 (1996).
34. M. Walther, D. G. Cooke, C. Sherstan, *et al.*, "Terahertz conductivity of thin gold films at the metal-insulator percolation transition," Phys. Rev. B **76**(12), 125408 (2007).
35. J. Horng, C.-F. Chen, B. Geng, *et al.*, "Drude conductivity of dirac fermions in graphene," arXiv (2010).
36. R. E. Glover and M. Tinkham, "Conductivity of superconducting films for photon energies between 0.3 and $40kT_c$," Phys. Rev. **108**(2), 243–256 (1957).
37. M. Koumans, D. Meulendijks, H. Middeljans, *et al.*, "Physics-assisted machine learning for thz time-domain spectroscopy: sensing leaf wetness," Sci. Rep. **14**(1), 7034 (2024).
38. A. Ghorbani, A. S. Nooramin, and A. Abdolali, "A general study for the complex refractive index extraction including noise effect using a machine learning-aided method," IEEE Access **12**, 11125–11134 (2023).
39. J. Dai, J. Zhang, W. Zhang, *et al.*, "Terahertz time-domain spectroscopy characterization of the far-infrared absorption and index of refraction of high-resistivity, float-zone silicon," J. Opt. Soc. Am. B **21**(7), 1379–1386 (2004).
40. J. Wang, C. Lan, C. Liu, *et al.*, "Generalizing to unseen domains: A survey on domain generalization," IEEE Trans. Knowl. Data Eng. **35**(8), 8052–8072 (2023).
41. C. Xu, Y. Cheng, J. Yu, *et al.*, "Improving graph out-of-distribution generalization on real-world data," arXiv (2024).
42. S. Minami, Y. Hayashi, S. Wu, *et al.*, "Scaling law of sim2real transfer learning in expanding computational materials databases for real-world predictions," arXiv (2024).
43. H. Zhang, L. Zhang, and Y. Jiang, "Overfitting and underfitting analysis for deep learning based end-to-end communication systems," in *11th International Conference on Wireless Communications and Signal Processing* (2019), pp. 1–6.
44. G. D. Barmparis and G. P. Tsironis, "Physics-informed machine learning for the covid-19 pandemic: Adherence to social distancing and short-term predictions for eight countries," Quant. Biol. **10**(2), 139–149 (2022).
45. B. Beddoes, N. Klokkou, J. Gorecki, *et al.*, "Dataset for "THz-TDS: extracting complex conductivity of two-dimensional materials via neural networks trained on synthetic and experimental data"," University of Southampton (2025), https://doi.org/10.5258/SOTON/D3433.