



Reconsidering Uncertainty from Frequency Domain Thermoreflectance Measurement and Novel Data Analysis by Deep Learning

Wenqing Shen, Diego Vaca & Satish Kumar


To cite this article: Wenqing Shen, Diego Vaca & Satish Kumar (2020) Reconsidering Uncertainty from Frequency Domain Thermoreflectance Measurement and Novel Data Analysis by Deep Learning, Nanoscale and Microscale Thermophysical Engineering, 24:3-4, 138-149, DOI: [10.1080/15567265.2020.1807662](https://doi.org/10.1080/15567265.2020.1807662)


To link to this article: <https://doi.org/10.1080/15567265.2020.1807662>

 View supplementary material [↗](#)

 Published online: 19 Aug 2020.

 Submit your article to this journal [↗](#)

 Article views: 182

 View related articles [↗](#)

 View Crossmark data [↗](#)

 Citing articles: 1 View citing articles [↗](#)



Reconsidering Uncertainty from Frequency Domain Thermoreflectance Measurement and Novel Data Analysis by Deep Learning

Wenqing Shen , Diego Vaca, and Satish Kumar

G.W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA

ABSTRACT

Frequency-domain thermoreflectance (FDTR) is a popular technique to investigate thermal properties of bulk and thin film materials. The FDTR data analysis involves fitting experimental data to a theoretical model whose accuracy may be affected by improper fitting approach and by convergence to local minima. This work proposes a novel data analysis approach using deep learning techniques. The developed deep learning model for FDTR (DL-FDTR) can accurately predict thermal conductivity, volumetric heat capacity and thermal boundary conductance with mean error below 5% for bulk samples coated with Au. DL-FDTR predictions can serve as an initial guess to the traditional fitting algorithms and can efficiently avoid local minima with regular fitting options, therefore improving the accuracy of data fitting and uncertainty evaluation.

ARTICLE HISTORY




Received 11 March 2020
Accepted 3 August 2020


KEYWORDS

FDTR; Deep Learning; Data Analysis; Uncertainty; Thermal Property

Introduction

In contrast to investigations of large bulk samples, investigating the thermal properties of nanoscale samples needs special techniques such as the 3ω method, time-domain thermoreflectance (TDTR) and frequency-domain thermoreflectance (FDTR) [1]. These techniques have different strengths and drawbacks. The 3ω method needs fabrication of a metal heater on the samples, while TDTR and FDTR are based on thermoreflectance and do not need complicated sample fabrication. The 3ω method is good at measuring cross-plane thermal conductivity but needs a narrower heater for the measurement of in-plane thermal conductivity of thin films [2]. The TDTR and FDTR methods have been widely used to study thermal conductivity, heat capacity, and thermal boundary conductance (TBC) of bulk materials [3–6] and thin films [6–8]. The TDTR and FDTR systems use a pump laser to heat the sample and a probe laser to detect the temperature response. The phase lag data between the pump and probe signals is fitted to a theoretical model to extract the unknown properties. In order to improve the signal to noise ratio, TDTR and FDTR usually require coating samples with a metal transducer of high thermoreflectance coefficient. TDTR uses a mechanical delay to tune the time delay between pump and probe signal, and thermal properties are extracted from the phase lag at different time delays. FDTR uses a modulated pump laser and extracts thermal properties from the phase lag using a range of modulation frequencies. TDTR with femtosecond pulsed laser might be better at the analysis of short time-scale non-equilibrium physics. FDTR, on the other hand, avoids the complex mechanical delay and could use less costly continuous lasers. Since FDTR only sweeps the modulation frequency, which can be easily tuned by a signal generator for pump modulation, its measurements can be obtained more quickly than those made using TDTR.

CONTACT Wenqing Shen  w.shen@gatech.edu; Satish Kumar  satish.kumar@me.gatech.edu  G.W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332

 Supplemental data for this article can be accessed [here](#).

FDTR and TDTR extract thermal properties by fitting measured data to a theoretical model, usually by least square fitting. The least square fitting task is in the form of $\text{argmin}_{\beta} |f(x, \beta) - y|^2$, where β is a vector of unknown parameters, x is known parameters, f is the theoretical model function, and y is the measured data. Nonlinear data-fitting tasks can be solved by the Levenberg-Marquardt [9] and trust-region [10] methods. These traditional methods require a starting point and converge to the solution after several iterations. While obtaining a global minimum is desired in all these cases, these algorithms may fall into local minima if the starting point is not good enough.

Measurement uncertainty indicates the confidence level in the results, and it is tightly related with the sensitivity of the parameters. The sensitivity of measurement for both FDTR and TDTR depends on the range of parameters, such as modulation frequency, beam spot size, transducer thickness, and the properties of the material to be measured. Low sensitivity of a fitting parameter can cause high uncertainty in the estimation of that parameter. For example, to measure the through-plane thermal conductivity of a thin film sample on low thermal conductivity substrates, the sample usually needs to be thicker than the thermal penetration depth and the modulation frequency of the pump laser needs to be chosen accordingly [11]. In-plane thermal conductivity measurement requires the laser spot size to be similar to or smaller than in-plane thermal penetration depth [11]. When fitting multiple parameters together, strong correlation between the parameters also decreases the fitting sensitivity [12].

The uncertainty of FDTR and TDTR measurements can be estimated through analytical methods [12, 13] and Monte Carlo simulations [8, 12]. Analytical methods accumulate uncertainties from the parameters and measurements. The quality of analytical estimation depends on the assumption behind the method. The analytical model for FDTR can be in good agreement with Monte Carlo simulations for tested cases in Ref [12], while the analytical model for TDTR without considering the correlation between parameters may overestimate the uncertainty compared with Monte Carlo results when the sensitivity is low to unknown properties for the investigated cases [8]. The analytical method is fast for the calculations but requires the proper values of parameters to be measured. Monte Carlo simulation is robust for the unknown parameters, but the computation cost is high, and it requires a robust fitting approach. Without a proper initial guess, Monte Carlo results may not be accurate.

Samples having different thermal properties may produce results with different phase lags between pump and probe signals if the phase is sensitive to a corresponding unknown parameter. Deep learning methods can serve as efficient tools to learn the pattern behind the data, and have demonstrated successes in many fields, such as object recognition [14, 15], and natural language processing [16]. To date, however, no study has been performed on deep learning application to FDTR data analysis, which will be the focus of the present work. This study first checks the uncertainty caused by local minima on the analysis of data measured by FDTR, and then demonstrates how a simple deep learning model can help in the prediction of thermal conductivity, heat capacity, and thermal boundary conductance while analyzing FDTR data. The deep learning model increased the fitting quality and uncertainty accuracy from Monte Carlo simulations by providing proper initial guess for data fitting.

Uncertainty caused by local minima while data-fitting

Data fitting seeks to minimize the residual and to find the global minimum point. Local minima are defined as points in the solution space that have a smaller residual than their near neighbors. For a 3-dimensional solution space, each point not at the boundary has 26 neighbors. The global minimum is the local minimum with the smallest residual among all the local minima. Depending on the fitting algorithms and initial guesses, the fitting may converge to any local minima. To check the effect of local minima, a discrete solution space is used. Phase lag data between the pump and probe signals is calculated for each sample by a theoretical model and is taken as a substitute for the experiment data. Each point in the solution space corresponds to its own phase lag data. For a sample, possible local minima resulting from the data fitting are searched by iterating through all points in the space and

comparing the norm of phase differences. For a typical bulk sample with a thin metal transducer, we consider 3 fitting parameters: the thermal conductivity (k) and volumetric heat capacity (C) of the bulk material, and the TBC (G) between the metal transducer and the bulk material. All the samples are assumed to be coated with 100 nm Au with thermal conductivity of 220 W/mK and volumetric heat capacity of 2.49 MJ/m³ K. The solution space has 20 points in log space between 0.1 and 10 for k (in W/mK), 100 points in linear space between 11 and 180 (in W/mK) for k , 100 points in linear space between 1.4 and 3.2 (in MJ/m³ K) for C , and 100 points in linear space between 20 and 150 (in MW/m² K) for G . The spot radius is 6.79 and 1.70 μm for pump and probe beam, respectively. The pump modulation frequency sweeps 50 values between 0.1 MHz and 10 MHz in the log space.

The average max error from local minima is 19.14% for the thermal conductivity, 12.15% for the heat capacity, and 25.57% for the TBC. 10,000 samples are sampled from the solution space. For each sample, all local minima are found and the maximum percentage error from local minima for each parameter is calculated. Here the global minimum is already known, since the phase data is calculated based on the properties of the sample. Usually, there is only one global minimum for one set of phase data, especially when the space is discrete rather than continuous. The distribution of error caused by local minima is plotted in Figure 1. Note that this error is in addition to the error caused by the phase measurement and uncertainty in the controlled parameters. For FDTR data analysis, if the fitting algorithm is not robust and the initial guess is far from the true value, the fitting error can be much higher. On the other hand, if the initial guess can be determined within a narrow range, the issues caused by local minimum may not be important.

Deep learning for FDTR

Deep learning can be used to predict unknown properties from experimental data without requiring an initial guess. The architecture of the feedforward deep learning model (DL-FDTR) used in the current analysis is shown in Figure 2. The model input is a series of phase lags at various modulation frequencies, and the output is G , C , or k , depending on the prediction goal. The model can be trained for a specific system configuration, such as laser spot size, transducer properties, etc. Trained models can then take the phase lag data of an unknown material measured with the same configuration, and predict G , C , or k for the new sample. One model predicts one unknown material property, and three models are needed to predict G , C , and k . Like traditional multi-parameter fitting, none of G , C , or k are included as the known values to any DL-FDTR model, e.g., C or k is not known to DL-FDTR model for G . A similar architecture can also output a vector and predict $G/C/k$ all at once, although that is not the approach in this study. All deep learning models are built and trained using Keras. The phase lag data is first normalized and then transformed to a fully connected (FC) layer. The first FC layer has an output size of 200 and uses the hyperbolic tangent (tanh) activation function. The output of the first batch normalization (BN) layer is concatenated to the output of the second BN. The

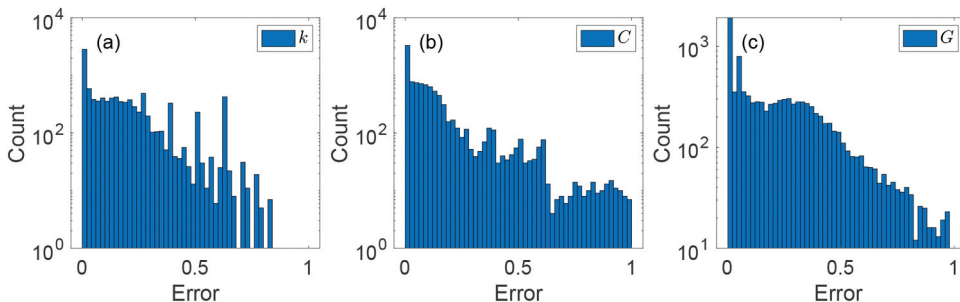


Figure 1. Error caused by local minima (a) thermal conductivity; (b) heat capacity; (c) thermal boundary conductance. The x-axis presents the fraction of error, e.g., error of 0.5 for variable R means $\Delta R/R = 0.5$.

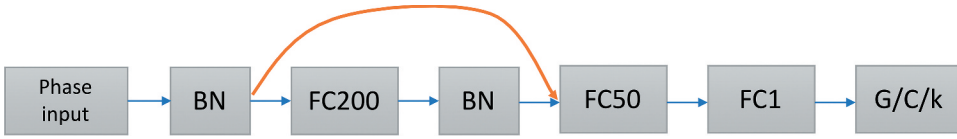


Figure 2. Architecture of deep learning model for DL-FDTR. FC200 has output size of 200 and uses tanh activation, FC50 has output size of 50 and uses ReLU activation. BN is batch normalization; FC1 is a fully connected dense layer whose output size is 1.

concatenation is inspired by DenseNets [15], which densely connects convolution layers. Adding the concatenation makes the training process more efficient. The second FC layer has output size of 50 and uses ReLU activation function. The last FC uses ReLU activation and outputs one value as the final prediction. The major layers are FC, which are also called dense layers that connects each input node to each output node linearly. The activation function introduces non-linearity to the output of FC and makes it possible to learn complicated information. The FC200 layer has output size larger than the input size of 50 and learns the combinational information from phase at each frequency. The FC50 and FC1 have output size smaller than input size and refine the characteristics of the input. The used architecture is not fully optimized but achieved sufficiently good performance. The training process uses the mean absolute percentage error as the loss function, which is consistent with the prediction goal.

Here the bulk material and thin film material are separately investigated. Training a deep learning model requires many samples, as the model learns from the statistics. To simplify the task, some assumptions are made: 1) all samples are coated by 100 nm Au with thermal conductivity of 220 W/mK and volumetric heat capacity of 2.49 MJ/m³ K; 2) the beam spot radius is 6.79 and 1.70 μm for pump and probe, respectively; 3) the pump modulation frequency sweeps 50 values between 0.1 MHz and 10 MHz in the log space; 4) the thin film thickness is 300 nm and a Si substrate is used for all thin film samples. The sample dataset for the bulk material is the same as the solution space used in the previous Section. There are three variables k , C , and G for the bulk material. Thin film samples have four variables G_1 (TBC between Au and thin film), G_2 (TBC between substrate and thin film), C and k . Here the C and k are for the material whose properties need to be measured/predicted. The dataset for the thin film samples includes 10 points for G_2 in the linear space between 20 and 150 (MW/m² K). The dataset size is 12,000,000 for the bulk material and 120,000,000 for the thin film material. Models with the same architecture are separately trained to predict $G/G_1/G_2$, C or k .

The developed DL-FDTR models can properly predict k , C , and G for both bulk and thin film materials. The percentage error distribution of DL-FDTR prediction for the bulk material is shown in Figure 3. The mean and standard deviation of error are listed in Table 1. The mean of absolute percentage error for all three parameters is less than 5%; this is much smaller than the error caused by

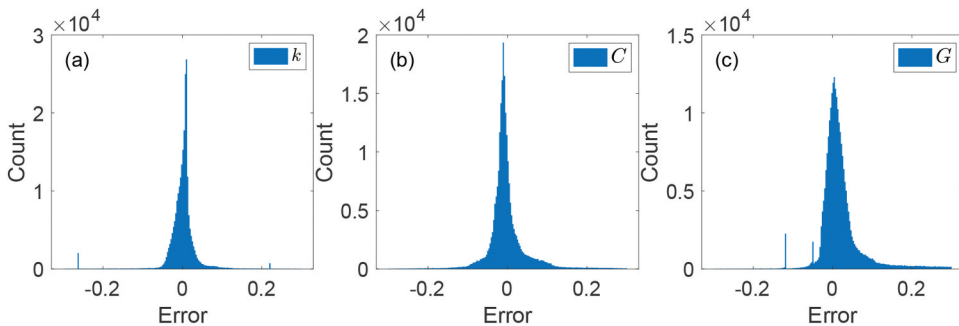


Figure 3. Prediction error distribution for bulk material (a) k , (b) C , (c) G . The x-axis presents the fraction of error ($\Delta R/R$).

Table 1. Statistics of prediction error for bulk material.

Parameter	DL-FDTR			Mean of absolute max error by local minima
	Mean of absolute error	Mean of error	Std of error	
<i>k</i>	2.34%	0.11%	4.96%	19.14%
<i>C</i>	3.23%	−0.13%	5.61%	12.15%
<i>G</i>	3.81%	2.35%	6.82%	25.57%

Table 2. Statistics of prediction error for thin film material.

Parameter	Mean of absolute error	Mean of error	Std of error
<i>k</i>	11.73%	6.91%	13.84%
<i>C</i>	3.77%	0.35%	5.54%
<i>G</i> ₁	6.82%	4.86%	9.04%
<i>G</i> ₂	9.35%	5.69%	19.62%

local minima. Indistinguishable phase data caused by the strong correlation between material properties and low sensitivity can result in high prediction error. The results for thin film material are shown in Table 2. The mean absolute error for *k*, *C*, and *G*₂ is below 12%. The mean absolute error for *k* and *G*₂ is much higher than that of *C* and *G*₁. The prediction performance for the thin films while considering four parameters is not as good as the performance for the bulk samples. Predicting four unknown parameters is more challenging than predicting three parameters because there is higher chance of strong correlation between the unknown parameters. Lower sensitivity could be another reason for higher average prediction error. If *C* is already known for the thin film, predicting for *k*, *G*₁ and *G*₂ will be easier than predicting all the 4 known parameters. Assuming $C = 1.64 \text{ MJ/m}^3\text{K}^{-1}$, and training the model on the subset of the dataset, the average error for *k* is 7.14%, which is much lower than the case with four unknown parameters, as shown in Table 3 and Figure 4. The prediction performance of *G*₁ and *G*₂ does not change significantly. Based on these results, the developed DL-FDTR model can either be used to roughly estimate thermal properties, or it can serve as a tool to choose the proper starting point for traditional data fitting. Note that using a pre-trained model for prediction takes negligible time.

Application of DL-FDTR for Monte Carlo uncertainty analysis

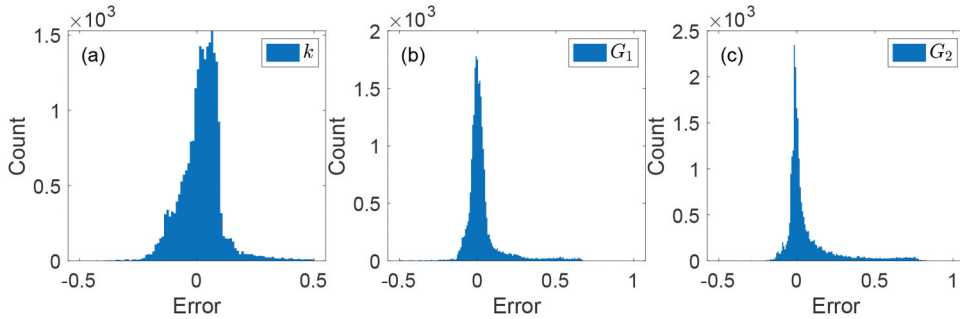
The uncertainty of the FDTR system originates from the phase signal measurement acquired by the lock-in amplifier, parameters for fitting, and fitting accuracy. The phase signal uncertainty is an overall outcome of fluctuations from electrical components, such as the amplifier, and from the optical system, such as laser power. Model parameter uncertainty is from the input parameters for fitting. Extracting the properties of unknown material requires fitting the phase data with a theoretical model, which includes controlled parameters that are assumed to be known. For example, when fitting thermal conductivity, volumetric heat capacity, and thermal boundary conductance for a bulk material, the controlled parameters include beam spot size, transducer's thermal properties and transducer's thickness. These controlled parameters can be collected from previous studies or separate measurements, which may not be 100% correct. Also, as the fitting might converge to local minima, the fitting itself might generate some uncertainty. The fitting could be challenging for parameters with high correlation or low sensitivity.

Uncertainty in estimating a parameter from FDTR data can be calculated through an analytical method. For FDTR measurement with least square fitting, the variance-covariance matrix of unknown parameters β_U is given by [12]

$$\text{Var}[\beta_U] = \left(J'_U J_U \right)^{-1} J'_U \left(\text{Var}[\Phi] + J_C \text{Var}[\beta_C] J'_C \right) J_U \left(J'_U J_U \right)^{-1} \quad (1)$$

Table 3. Statistics of prediction error for thin film material with known C .

Parameter	Mean of absolute error	Mean of error	Std of error
k	7.14%	2.48%	10.41%
G_1	7.07%	2.18%	14.76%
G_2	9.87%	6.39%	17.9%


Figure 4. Prediction error distribution for a thin film with $C = 1.64\text{MJm}^{-3}\text{K}^{-1}$ (a) k , (b) G_1 , (c) G_2 . The x-axis presents the fraction of error ($\Delta R/R$).

where J is the Jacobian matrix, U represents the unknown parameters, C represents the controlled parameters, Φ is the measured phase lag. The diagonal part of the variance-covariance matrix $\text{Var}[\beta_U]$ is the uncertainty of the unknown parameters caused by the measurement uncertainty $\text{Var}[\Phi]$ and the uncertainty of the controlled parameters $\text{Var}[\beta_C]$. The non-diagonal part of $\text{Var}[\beta_U]$ is the covariance between parameters, which could be used to check the linear dependence and its effect on the confidence intervals.

Equation (1) provides a good estimation of uncertainty if the global optimum of β_U is already found. The uncertainty calculation could be very poor if the used solution is not close to the global minimum. For an unknown material, and without knowing the correct value, one can get better fitting results by using multiple starting points and picking the result with the least residual. However, it is difficult to verify if a solution is the global minimum, as the unknown parameters are in n -dimension continuous space, where n is the number of parameters to be fitted. If the parameters used for the analytical uncertainty calculation is not close to the true properties, the Jacobian matrix calculation and the resulting uncertainty can be affected. In addition, the analytical method assumes the gradient is constant for variant parameters in certain windows. The error caused by this assumption might be negligible if the uncertainty of the controlled parameters is low, but may be high if the uncertainty of the controlled parameters is high or if the sensitivity to the controlled parameters is high.

Unlike analytical methods that only consider one possible value set of parameters, Monte Carlo simulation considers a large enough number of value sets to get the unknown parameter distribution. Monte Carlo simulation for FDTR usually assumes normal distributions for the controlled parameters and phase measurement. In each iteration, controlled parameters and phase data are randomly sampled from the corresponding distributions, then fitted with a theoretical model to generate a value set for the unknown parameters. After enough iterations, the mean value and associated uncertainty for the unknown parameters can be extracted from the distribution of possible values. As

each Monte Carlo iteration includes fitting, a good initial guess is necessary in order to avoid high fitting uncertainty.

DL-FDTR can increase the accuracy in evaluating mean value and uncertainty from Monte Carlo simulations. Here, the role of DL-FDTR is to provide a good initial guess for the given parameters and phase dataset. Here we compare the results from three different Monte Carlo approaches, TR, LM and DL-LM. TR uses a trust-region reflective fitting algorithm with advanced fitting options provided by the Matlab lsqnonlin module. TR is reliable according to our test and serves as a baseline, although it still has chances of converging to local minima. LM uses the Levenberg Marquardt fitting algorithm through Matlab lsqnonlin and uses default options. DL-LM uses the same fitting algorithm as LM and takes the DL-FDTR prediction as an initial guess. A fair initial guess is required for TR and LM to make the comparison meaningful. The initial guess of TR is based on the actual properties which are used to generate the raw phase data without additional noise, and is offset randomly, assuming the offset follows a normal distribution whose standard deviation equals 20%. The initial guess of LM is based on the fitting results from TR and is offset in the same way as to pick initial guess for TR.

Verification of the Monte Carlo simulation was done by performing simulations for a SiO₂ sample similar to Ref [12]. The controlled parameters are the same as Ref [12]. The pump and probe radius are $2.8 \pm 0.03 \mu\text{m}$ and $2.3 \pm 0.02 \mu\text{m}$, respectively. The Au thickness is $81 \pm 1.5 \text{ nm}$, TBC is 51 MW/m²K. 1000 phase lag datasets are created by adding random noise to the ideal data. The fitting uses the TR algorithm and takes a slightly offset mean value as an initial guess. At the phase noise level of 0.1 degree, uncertainty in estimating thermal conductivity, heat capacity, and thermal boundary conductance are 8.66%, 3.24%, and 16.97%, respectively, compared with 8.6%, 3.4%, and 14.6% from Monte Carlo results in Ref [12]. The small difference may be caused by different approaches to consider the phase noise level. Here we add random phase noise to the calculated phase data according to the noise level, and the generated phase data may not be exactly the same as in Ref [12]. The uncertainty values are in good agreement with analytical values in this scenario.

Monte Carlo simulations are tested for 3 different materials – SiO₂, Si, and sapphire – with different values of TBC. Sample properties listed in Table 4 are used to generate ideal data, and then phase noise is manually added. The materials cover both low and high thermal conductivities. 100 nm gold is used as the transducer. The modulation frequency series includes 50 points in the logarithm space of 0.1–10 MHz. The other parameters are the same as in the last Section. To prove the effectiveness of applying DL-FDTR to data analysis, we consider the uncertainty of the phase measurement and transducer thickness. The standard deviation of gold thickness is assumed to be 1.5 nm. Unless specified, the standard deviation in phase noise is 0.1 degree. The DL-FDTR model is separately trained for different Au thicknesses ranging from 96 to 104 nm. The prediction error of the deep learning model is listed in Table 5. The mean absolute error of the deep learning model is below 2.6%, 4.3% and 4.3% for k , C , and G , respectively. In the Monte Carlo simulations, the Au thickness sampled from the normal distribution is rounded to the closest thickness value in Table 5.

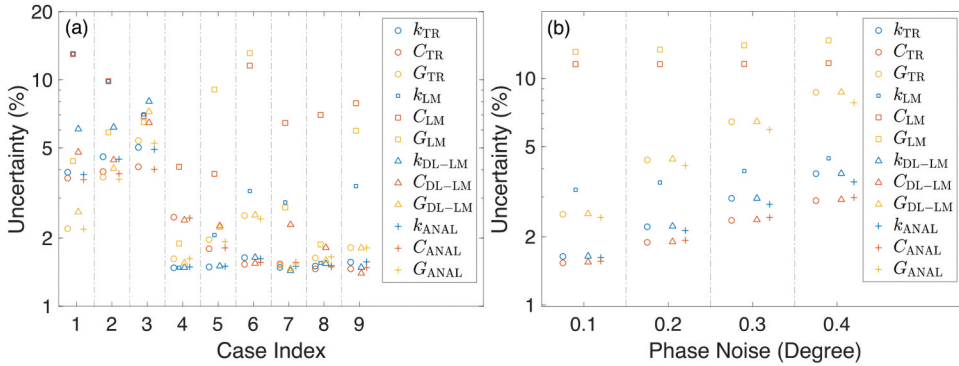
The uncertainty in k , C , and G for the cases 1–9 listed in Table 4 are shown in Figure 5(a). Analytical calculations based on Eq. (1) are also listed for comparison. The effect of phase noise level is considered for the same material as in case 6, and the results are shown in Figure 5(b). The analytical calculations could provide good uncertainty estimation for the least square fitting, given that we know the mean true value of fitting parameters. The uncertainties calculated from Monte Carlo simulations using TR are consistent with the analytical calculations for all the test cases. The uncertainties from LM are greater than analytical values for the majority of tests, while the

Table 4. Sample properties for Monte Carlo simulations.

Parameter	$k(\text{Wm}^{-1}\text{K}^{-1})$	$C(\text{MJm}^{-3}\text{K}^{-1})$	$G(\text{MWm}^{-2}\text{K}^{-1})$
SiO ₂	1.4	1.63	25, 50, 75 for case 1, 2, 3
Si	140	1.65	25, 50, 75 for case 4, 5, 6
Sapphire	33	3.00	25, 50, 75 for case 7, 8, 9

Table 5. Prediction error of DL-FDTR model for different Au thickness.

Au thickness (nm)	Mean abs error for k (%)	Std of error for k (%)	Mean abs error for C (%)	Std of error for C (%)	Mean abs error for G (%)	Std of error for G (%)
96	2.54	4.95	3.83	6.45	4.20	6.35
97	2.48	4.97	3.69	6.26	3.58	6.09
98	2.44	5.44	3.64	6.14	3.56	5.95
99	2.29	5.14	3.37	5.81	3.73	6.38
100	2.34	4.96	3.23	5.61	3.81	6.82
101	2.41	5.04	4.23	7.13	3.09	5.68
102	2.53	5.28	3.64	6.07	2.86	6.05
103	2.42	5.11	3.56	6.28	3.30	6.87
104	2.27	4.97	4.22	6.84	3.29	6.62


Figure 5. Uncertainty in k , C and G from Monte Carlo simulations (TR, LM, DL-LM) and analytical calculations. (a) Compare the uncertainties for different samples; (b) Compare the uncertainties for 4 different phase noise levels. The sample is the same as case 6.

uncertainties from DL-LM are similar to analytical values for most of the tests. As there is no absolute way to tell whether a fitting result is a local minimum or a global minimum, the fitting residual magnitude is used to indicate the fitting quality. Upon checking the local minima of the ideal phase data without phase noise, the residual of local minima with longer distance from the true value is usually larger than the residual of local minima with shorter distance from the true value. TR results in the smallest average fitting residual among the three fitting approaches. DL-LM has lower residual than LM for most of the test cases. Based on the analytical uncertainty calculation and fitting residual, LM tends to converge into fitting results far from the true value, therefore overestimating the uncertainties. Figure 6 shows the Monte Carlo simulation results for case 6 with phase noise of 0.4 degree. The distributions from TR and DL-LM are very similar and are close to the normal distribution. The distribution from LM shows obvious local minima for C and G , causing their uncertainty to be overestimated. By taking DL-FDTR prediction as the initial guess, DL-LM eases the issue of overestimating uncertainty caused by convergence to local minima, which is observed in the standalone LM approach. Note that the standalone LM takes slightly offset TR fitting results as the initial guess, which is fair to guess as the initial points. DL-FDTR might predict better than TR offset, which used a normal distribution with $\text{std} = 20\%$ and provide a better initial guess for nonlinear least square fitting.

Demonstration of DL-FDTR to real measurement

DL-FDTR is further applied and validated on real measurements. The experiment uses a Si sample coated by 101 ± 1.5 nm Au as transducer and 1 nm Ti as adhesive layer. The uncertainty in Au thickness is obtained from AFM measurements. The phase signals are detected by a lock-in amplifier with 3.4 Hz bandwidth with modulation frequency in the range of 100 kHz – 4 MHz. Figure 7 shows

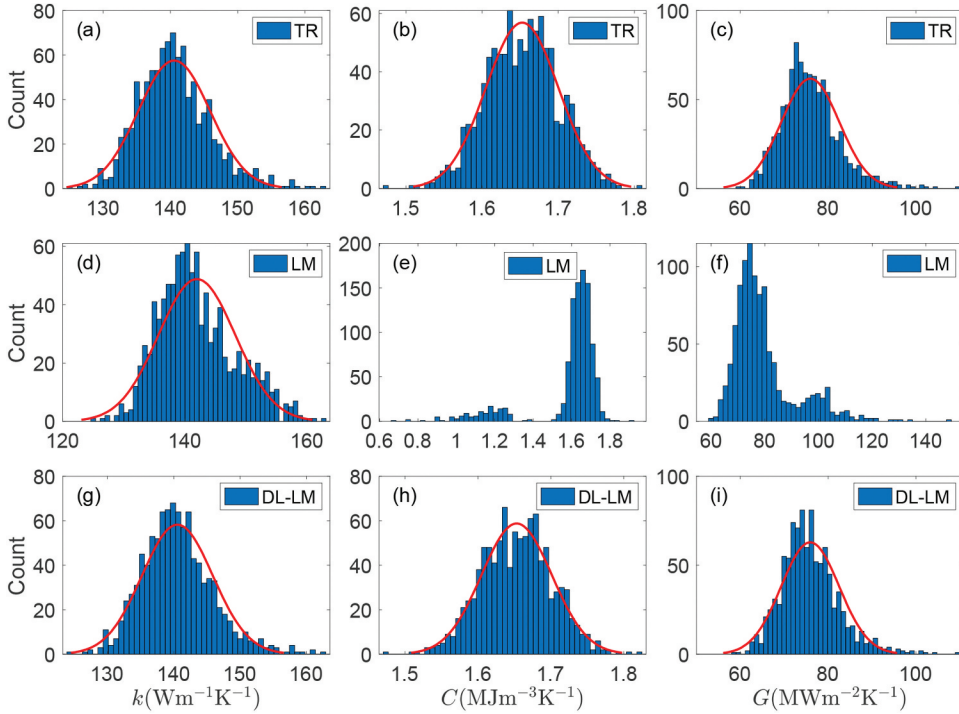


Figure 6. Histogram of fitting results from TR, LM, DL-LM. The red curve shows the fitting with normal distribution.

the averaged phase signal and best fitting with 160 kHz – 4 MHz frequency. Sensitivity at 100 kHz – 160 kHz is low, and the data at that window is hard to fit. The following discussion is based on the data in the frequency range of 160 kHz – 4 MHz. Phase noise is extracted from the standard deviation of 1000 measurements at each modulation frequency, and the RMS of phase noise is 0.14 degree. Monte Carlo simulations similar to those presented in the previous Section are performed and compared with the analytical calculations. Although uncertainty in Au thermal properties and beam spot size also introduce certain amount of uncertainty to the final results, these are not considered in this demonstration, because considering those may require including additional deep learning models. Future studies may include designing a general deep learning model that also considers controlled parameters of the thermal model as inputs, so that only one deep learning model is needed for various thermal parameters of the model.

Figure 8 shows the distribution of fitted parameters using a different fitting approach. Table 6 lists details of mean values and uncertainties according to Monte Carlo simulations and analytical calculations. TR and DL-LM have similar distributions for all three parameters, which are close to the normal distributions. LM has a different mean value and higher uncertainty than TR and DL-LM. The distribution from LM is likely to be a mixture of two normal distributions. The major distributions on the right side of Figure 8(d-f) are at values similar to TR and DL-LM, while the secondary distributions are at lower parameter values. Examples of fitting to the two bumps are shown in Figure S1 (see supplemental information). Table 6 lists the mean squared norm of residual (ResNorm) from each fitting method. TR has the lowest ResNorm, meaning the overall fitting quality is the best among these three fitting methods. The ResNorm of DL-LM is slightly higher than TR, but the difference is negligible. The ResNorm of LM is more than twice of that of TR, which is caused by local minima with values located at the secondary distribution in Figure 8(d-f).

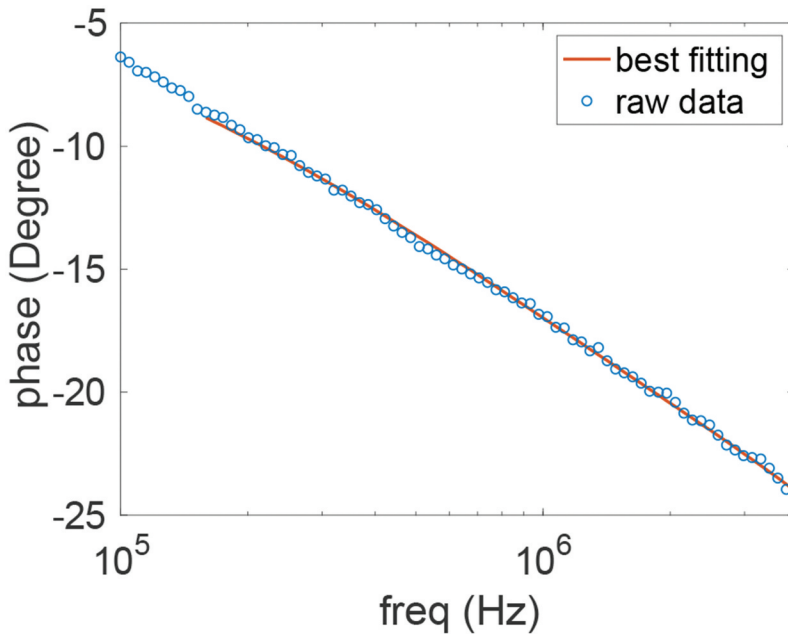


Figure 7. Phase signal and best fitting for Au/Si data. Phase data for modulation frequencies in the range of 160 kHz – 4 MHz is used for the fitting.

The analytical uncertainty listed in Table 6 is calculated using the mean value from TR, assuming TR converged to the value equal or close to the global minima. The assumption is valid for the studied case as the TR with different options used in this research usually converges to a same narrow range for this specific case, even with a very bad initial guess, according to the additional tests. Note that TR does not guarantee finding the global minimum. Although TR works well for this case, it may not work well for the other cases. The uncertainty from TR and DL-LM is close to the analytical results, while uncertainty from LM is overestimated. DL predicts reasonably well for k and C , but is slightly off for G . The ResNorm of DL is calculated by plugging the predicted parameters into the theoretical model and comparing the difference between the simulated phase data and the experimental data. Although the ResNorm of DL is even higher than that of LM, using DL-FDTR prediction as initial guess can efficiently help avoid local minima when using LM algorithm.

Discussion

By using DL-FDTR prediction as the initial guess for traditional fitting algorithms like Levenberg Marquardt and trust-region reflective, one can save effort in finding the best fitting options for each sample, and there will be no need to try various initial guesses in a large range. If the initial guess can be determined to be in a narrow range, the issues from fitting itself will not be as crucial as the issues arising from the properties which are totally unknown, and in this case traditional data fitting could be good enough. The issue of local minima can be avoided if there are fewer unknown parameters, as more unknowns results in more non-linearity. Note that the DL-FDTR predicts well when the phase data is unique to the certain properties, but may not distinguish one fitting result to another if the theoretical phase has subtle differences.

The computational time taken by DL-FDTR compared with the traditional fitting algorithms is negligible. Considering an example of Monte Carlo simulations for case 6 with phase noise of 0.4 degree, DL-FDTR took 1.97 ms, TR took 145.29 ms, DL-LM took 204.63 ms (including DL

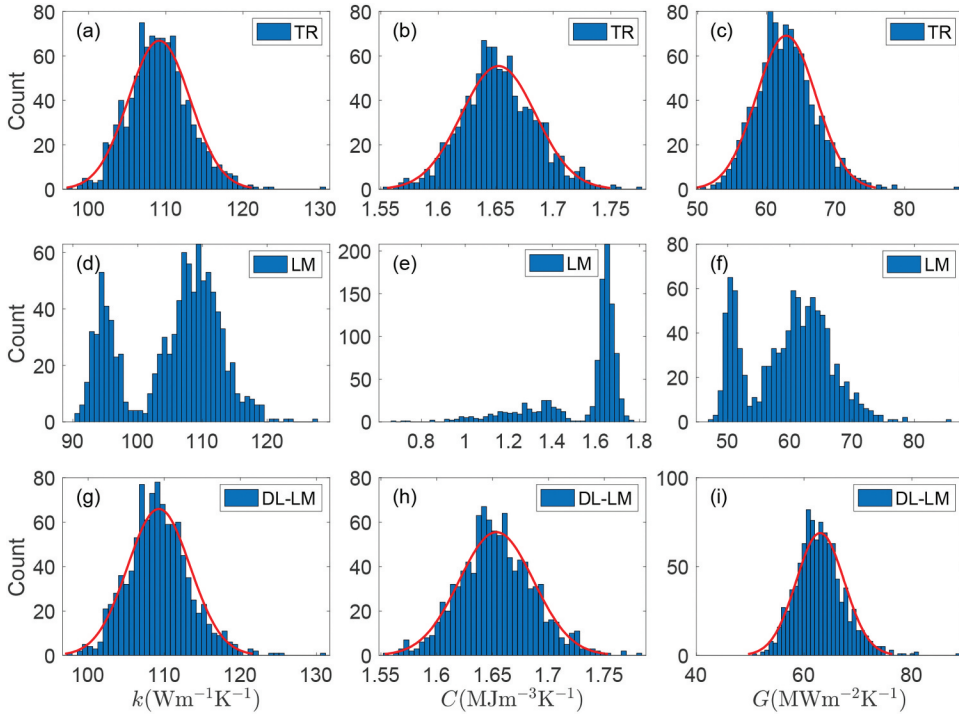


Figure 8. Histogram of the fitting results from TR, LM, DL-LM for Au/Si sample. The red curve shows the fitting with normal distribution.

Table 6. Fitting results of Au/Si from Monte Carlo simulations. Uncertainty value is listed in bracket. ResNorm is the squared norm of the residual.

Method	$k(\text{Wm}^{-1}\text{K}^{-1})$	$C(\text{MJm}^{-3}\text{K}^{-1})$	$G(\text{MWm}^{-2}\text{K}^{-1})$	ResNorm
TR	109.16 (3.66%)	1.652 (1.96%)	62.83 (6.89%)	3.1719
LM	105.32 (6.96%)	1.546 (12.60%)	59.77 (10.85%)	8.0128
DL-LM	109.29 (3.76%)	1.653 (2.01%)	62.98 (7.08%)	3.1730
DL	111.94 (5.84%)	1.664 (3.18%)	69.40 (10.07%)	13.0071
Analytical	(3.79%)	(2.06%)	(7.07%)	

prediction), and LM took 550.92 ms for fitting/prediction per dataset. The benchmark was run on a desktop with i7-4790 CPU and 16GB memory. With better initial guess from DL-FDTR, traditional fitting algorithms will take less iterations to converge as can be observed while using DL-LM.

Conclusion

Data analysis of this FDTR experiment may have additional uncertainty caused by local minima when fitting the experimental data to the theoretical model. This kind of uncertainty could be over 10% for thermal conductivity and volumetric heat capacity of bulk samples and thermal boundary conductance between metal transducer and bulk sample. The quality of traditional fitting algorithms depends on the initial guess. Uncertainty from Monte Carlo simulation may be overestimated if the fitting algorithm is not carefully tuned and if the initial guess is far from global minima. In this study, we proposed a novel data analysis approach, which combines deep learning and traditional fitting algorithms. The developed deep learning model can predict thermal conductivity, volumetric heat capacity, and thermal boundary conductance with an average error below 5% for bulk samples coated

with Au. This DL-FDTR model can also accurately predict thermal conductivity of thin film and two interfacial conductance for thin film samples, when volumetric heat capacity of thin film is given. Using DL-FDTR prediction as the initial guess for the traditional fitting algorithms, the problem of converging to local minima can be eased, and Monte Carlo simulations can result in more accurate uncertainty values. The proposed application of deep learning could be applicable to other data analysis tasks, which involve data fitting.

ORCID

Wenqing Shen  <http://orcid.org/0000-0002-1258-3579>

References

- [1] D. Zhao, X. Qian, X. Gu, S. A. Jajja, and R. Yang, "Measurement techniques for thermal conductivity and interfacial thermal conductance of bulk and thin film materials," *J. Electron. Packag.*, vol. 138, no. 4, 2016. DOI:10.1115/1.4034605
- [2] C. Dames, "Measuring the thermal conductivity of thin films: 3 Omega and related electrothermal methods," *Annual Rev. Heat Trans.*, vol. 16, no. 1, pp. 7–49, 2013. DOI: 10.1615/AnnualRevHeatTransfer.v16.20.
- [3] K. T. Regner, S. Majumdar, and J. A. Malen, "Instrumentation of broadband frequency domain thermoreflectance for measuring thermal conductivity accumulation functions," *Rev. Sci. Instrum.*, vol. 84, no. 6, pp. Art no. 064901, Jun. 2013. DOI: 10.1063/1.4808055.
- [4] A. J. Schmidt, R. Cheaito, and M. Chiesa, "A frequency-domain thermoreflectance method for the characterization of thermal properties," *Rev. Sci. Instrum.*, vol. 80, no. 9, pp. 094901, Jan 9. 2009. DOI: 10.1063/1.3212673.
- [5] J. Yang, C. Maragliano, and A. J. Schmidt, "Thermal property microscopy with frequency domain thermoreflectance," *Rev. Sci. Instrum.*, vol. 84, no. 10, pp. 104904, 2013. DOI: 10.1063/1.4824143.
- [6] J. Liu, *et al.*, "Simultaneous measurement of thermal conductivity and heat capacity of bulk and thin film materials using frequency-dependent transient thermoreflectance method," *Rev. Sci. Instrum.*, vol. 84, no. 3, pp. 034902, 2013. DOI: 10.1063/1.4797479.
- [7] D. B. Brown, *et al.*, "Spatial mapping of thermal boundary conductance at metal–molybdenum diselenide interfaces," *ACS Appl. Mater. Interfaces*, vol. 11, no. 15, pp. 14418–14426, 2019. DOI: 10.1021/acsami.8b22702.
- [8] T. L. Bougher, *et al.*, "Thermal boundary resistance in GaN films measured by time domain thermoreflectance with robust monte carlo uncertainty estimation," *Nanoscale Microscale Thermophys. Eng.*, vol. 20, no. 1, pp. 22–32, 2016. DOI: 10.1080/15567265.2016.1154630.
- [9] Wikipedia. "Levenberg–Marquardt algorithm." Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Levenberg%E2%80%93Marquardt_algorithm&oldid=930437222 (accessed 2019).
- [10] Wikipedia. "Trust region." Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Trust_region&oldid=911203117 (accessed 2019).
- [11] P. Jiang, X. Qian, and R. Yang, "Tutorial: time-domain thermoreflectance (TDTR) for thermal property characterization of bulk and thin film materials," *J. Appl. Phys.*, vol. 124, no. 16, pp. 161103, 2018. DOI: 10.1063/1.5046944.
- [12] J. Yang, E. Ziade, and A. J. Schmidt, "Uncertainty analysis of thermoreflectance measurements," *Rev. Sci. Instrum.*, vol. 87, no. 1, 2016. DOI:10.1063/1.4939671
- [13] C. Wei, X. Zheng, D. G. Cahill, and J.-C. Zhao, "Invited article: micron resolution spatially resolved measurement of heat capacity using dual-frequency time-domain thermoreflectance," *Rev. Sci. Instrum.*, vol. 84, no. 7, pp. 071301, Jan 7. 2013. DOI: 10.1063/1.4815867.
- [14] L. Liu *et al.*, "Deep learning for generic object detection: a survey," *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 261–318, 2019. DOI: 10.1007/s11263-019-01247-4.
- [15] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," In 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 21–26 July 2017, pp.2261–2269, doi:10.1109/CVPR35066.2017.
- [16] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing [Review article]," *IEEE Comput. Intell. Mag.*, vol. 13, no. 3, pp. 55–75, 2018. DOI: 10.1109/mci.2018.2840738.