

# SCIENTIFIC REPORTS



OPEN

## Prediction of thermal boundary resistance by the machine learning method

Received: 2 March 2017

Accepted: 26 June 2017

Published online: 2 August 2017

Tianzhuo Zhan, Lei Fang &amp; Yibin Xu

Thermal boundary resistance (TBR) is a key property for the thermal management of high power micro- and opto-electronic devices and for the development of high efficiency thermal barrier coatings and thermoelectric materials. Prediction of TBR is important for guiding the discovery of interfaces with very low or very high TBR. In this study, we report the prediction of TBR by the machine learning method. We trained machine learning models using the collected experimental TBR data as training data and materials properties that might affect TBR as descriptors. We found that the machine learning models have much better predictive accuracy than the commonly used acoustic mismatch model and diffuse mismatch model. Among the trained models, the Gaussian process regression and the support vector regression models have better predictive accuracy. Also, by comparing the prediction results using different descriptor sets, we found that the film thickness is an important descriptor in the prediction of TBR. These results indicate that machine learning is an accurate and cost-effective method for the prediction of TBR.

Tailoring the thermal resistance of materials is vital for the thermal management of high power micro- and opto-electronic devices and for the development of high efficiency thermal barrier coatings and thermoelectric materials<sup>1–3</sup>. The overall thermal resistance of a material system is composed of the thermal resistance of the constituent materials and the thermal boundary resistance (TBR) between those materials. Heat in dielectric materials and semiconductors is transported predominantly by phonons, which undergo scattering in materials by interacting with defects, other phonons, boundaries, isotope, etc<sup>4</sup>. These processes cause the thermal resistance of the constituent materials. When heat passes through an interface between two materials, a temperature discontinuity occurs at the interface. TBR is defined as the ratio of the temperature discontinuity at an interface to the heat flux flowing across that interface. Typically, the characteristic length scales of the nanostructured materials are from several to hundreds of nanometers. By comparison, the phonon mean free paths (MFPs) in crystalline materials can be micrometers long<sup>5</sup>. When the characteristic length scales of nanostructured materials are shorter than the phonon MFPs, phonon transport is ballistic instead of diffusive. Furthermore, the number of interfaces increases greatly with decreasing characteristic length scales in nanostructured materials, such as superlattices. Thus, TBR can dominate the thermal resistance of the constituent materials and become a major component of the overall thermal resistance in nanostructured materials<sup>1–3</sup>. TBR is attributed to the mismatch between the phonon spectra of the materials on both sides of the interface<sup>5</sup>; interfacial properties, such as interfacial roughness, interdiffusion, bonding strength, and interface chemistry, also have a considerable effect<sup>1–3</sup>. Finding interfaces with very low or very high TBR experimentally is a high cost and time-consuming approach. Therefore, in order to guide the discovery of such interfaces, a reliable method for the prediction of TBR is necessary.

Several methods have been used for predicting TBR. The most common methods are the acoustic mismatch model (AMM) and the diffuse mismatch model (DMM)<sup>6</sup>. In the AMM, phonons are treated as plane waves, and the materials in which the phonons propagate are treated as continua. The transmission probabilities of phonons are calculated from the acoustic impedances on each side of the interface. A crucial assumption made in the AMM is that no scattering occurs at the interface. The assumptions of wave nature of phonon transport and specular scattering at the interface make the AMM valid when predicting TBR at low temperatures and at ideal interfaces. By comparison, in the DMM, completely diffuse scattering at the interface is assumed: a scattered phonon has no memory of its modes (longitudinal or transverse) and where it came from. The transmission probabilities of phonons are determined by the mismatch of the phonon density of states (DOS) on each side of the interface.

National Institute for Materials Science, 1-2-1 Sengen, Tsukuba, 305-0047, Japan. Correspondence and requests for materials should be addressed to Y.X. (email: [XU.Yibin@nims.go.jp](mailto:XU.Yibin@nims.go.jp))

Another crucial assumption made in the DMM is that phonons are elastically scattered: the transmitted phonons have the same frequencies with the incident phonons. These assumptions make the DMM invalid for an interface where inelastic phonon scattering occurs. In recent years, the molecular dynamics (MD) simulation method has emerged as a method for predicting TBR<sup>7,8</sup>. In MD simulations, no assumptions concerning the nature of phonon scattering are required; the only required input to an MD simulation is the description of the atomic interactions. However, MD simulations also have limitations. The specification of the atomic interactions is typically done using empirical interatomic potential functions, and the results of MD simulations are realistic only when the specified atomic interactions mimic the forces experienced by the ‘real’ atoms<sup>9</sup>. Specifying an adequate description of atomic interactions that is applicable to various systems is a challenging task. Furthermore, MD simulations are computationally expensive and time-consuming, especially for simulation cells containing a large number of atoms. Thus, it is imperative to develop an accurate and cost-effective method to guide the discovery of interfaces with very low or very high TBR. In this work, we demonstrate a machine learning method for predicting TBR with much better predictive accuracy than the commonly used AMM and DMM.

## Methods

Machine learning is a subfield of computer science that allows computers to improve their performance through experience<sup>10</sup>. When using machine learning to make predictions, one should consider three key ingredients: training data, descriptors, and machine learning algorithms. In this work, the training data include a large amount of experimental TBR data collected from 62 published papers<sup>1–3, 11–69</sup>. The data consist of a total of 876 TBRs measured for 368 interfaces as a function of measurement temperature and other conditions. The 368 interfaces comprise 45 different materials. The descriptors are measurement temperature, film thickness, heat capacity, thermal conductivity, Debye temperature, melting point, density, speed of sound (longitudinal and transverse), elastic modulus, bulk modulus, thermal expansion coefficient, and unit cell volume; collectively, these form the set which we define as “all collected descriptors”. We chose these descriptors from two points of view. Firstly, the descriptors are the measurement conditions and the properties of the materials constituting the interfaces that might affect TBR. Secondly, the descriptors should be easily collected. Namely, the values of the descriptors are clearly given in the references. The data of the descriptors are collected from references including the Internet, published papers<sup>1–3, 11–69</sup>, and databases, such as Atom Work in the NIMS MatNavi database<sup>70</sup> and the TPRC data series<sup>71</sup>. The descriptor data for all the materials and their origins are shown in the supplementary files. The machine learning algorithms used in this work and the model fitting process are briefly introduced in the following paragraphs.

**Generalized linear regression (GLR).** Linear regression is a method describing a linear relationship between a response and one or more descriptors<sup>72</sup>. The simple linear regression model can be described as

$$E(y_i) = \mu_i = x_i^T \beta, \quad i = 1, \dots, N. \quad (1.1)$$

GLR extends simple linear regression. Unlike the simple linear regression, whose response has a normal distribution, the GLR response could be normal, binomial, Poisson, or something else. In GLR, a link function  $f$  is applied to the linear description, so the GLR model becomes

$$f(\mu_i) = f(E(y_i)) = x_i^T \beta. \quad (1.2)$$

In our TBR prediction analysis, we choose the logarithm as the link function,  $\log(E(y_i)) = x_i^T \beta$ , and the response has a Poisson distribution.

**Least-absolute shrinkage and selection operator regularization (LASSO-GLR).** LASSO regularization is a shrinkage method that imposes a penalty term to limit a model’s complexity<sup>73</sup>. It can identify important descriptors, select descriptors among redundant descriptors, and produce fewer coefficients in the model formula. Moreover, it can address descriptor multicollinearity. LASSO regularization in GLR is an extension of simple LASSO regularization. It is defined by

$$\min_{\beta_0, \beta} \left( \frac{1}{N} \text{Deviance}(\beta_0, \beta) + \lambda \sum_{j=1}^p |\beta_j| \right). \quad (1.3)$$

where the second term with coefficient  $\lambda$  is a penalty term to balance model accuracy and complexity, and *Deviance* is the deviance between the response and the model fit by  $\beta_0$  and  $\beta$  coefficients. In this study, the response has a Poisson distribution, and the link function is logarithm. The model is first given with all collected descriptors. After training, the number of coefficients in model is reduced due to the LASSO screening.

**Gaussian process regression (GPR).** The GPR model is a nonparametric probabilistic model<sup>74</sup>. The goal of the GPR is to find a probabilistic distribution of new output given the training data and new input data,  $P(y_{new} | y_{train}, x_{train}, x_{new})$ . Given that a latent variable  $f(x_i)$  is a Gaussian process, then the joint distribution of a finite number of latent variables is also a Gaussian process, defined normally by  $f(x) \sim GP(0, k(x, x'))$ , where  $k(x, x')$  is the kernel (covariance) function. In GPR, for known prior distributions of  $f$  and  $y$ , the posterior distribution of  $f_{new}$  can be computed. In our study, the kernel function  $k(x_i, x_j)$  is a radial basis function which is

$$k(x_i, x_j) = \sigma_f^2 \exp\left(-\frac{(x_i - x_j)^T(x_i - x_j)}{2\sigma_l^2}\right). \quad (1.4)$$

The predicted mean  $\bar{f}_{new}$  is given by

$$\bar{f}_{new} = K(x_{new}, X)[K(X, X) + \sigma^2 I]^{-1}y. \quad (1.5)$$

where  $\sigma^2$  is the noise variance.

**Support vector regression (SVR).** SVR is an effective method of machine learning for regression problems<sup>75</sup>. The aim of SVR is to find a function to map a non-linear input space into a high dimensional feature space, so linear regression can be applied in this space. In SVR, the  $\varepsilon$ -insensitive loss function is commonly used, and a penalty factor  $C$  is chosen to control the balance between the model complexity and training errors. Like GPR, a kernel function is used to transform the data. The SVR optimization problem can be solved by quadratic programming algorithms, and the function to predict new data can be made by

$$f(x) = \sum_{i=1}^N (\alpha_i - \alpha_i^*) K(x, x_i) + b. \quad (1.6)$$

where  $\alpha, \alpha^*$  are Lagrange multipliers, and  $K$  (kernel function) is a radial basis function, which in this study is

$$K(x_i, x_j) = \exp\left(-\frac{(x_i - x_j)^T(x_i - x_j)}{2\gamma^2}\right). \quad (1.7)$$

We next introduce the model fitting process. For the GLR model, we use all the data to fit the model. The  $\beta$  coefficients are calculated by the least squares approach. For the LASSO-GLR, the GPR and the SVR models, model fitting is performed by finding optimal parameters  $\lambda$  (LASSO-GLR),  $\sigma_f$  and  $\sigma_l$  (GPR),  $C$ ,  $\varepsilon$ , and  $\gamma$  (SVR). In this study, 5- and 10-fold cross-validations are performed to find these parameters. Since the results of the two cross-validations show very slight difference (<1%), only the results of the 10-fold cross-validation are shown in this study. The overall data are randomly partitioned into ten folds. Model training is done with nine folds and validation is done with the remaining one. This process is repeated ten times with each fold used exactly once as the validation data. Furthermore, to investigate the effect of random partition on the predictive accuracy, we try four different random partitions to split data by assigning different seeds to random generator. The calculated predictive accuracies of the four partitions show slight differences (<1%). Therefore, we consider the model's predictive accuracy difference caused by the random data choice is very small. In this study, the average predictive accuracies of the four random partitions are used.

In our study, the model fitting is performed with MATLAB statistical software<sup>76</sup> under different initial settings based on the machine learning algorithm, training data and descriptor set.

## Results and Discussion

We have trained four models to predict TBR, which correspond to the four machine learning algorithms described above. We first predict the TBR of all the interfaces using the AMM and DMM. The details of the AMM and DMM predictions are described in ref. 5. For the AMM and DMM, the TBR can be written as

$$\frac{1}{R} = \frac{1}{2} \sum_j c_{i,j} \int_0^{\pi/2} \alpha_{i,j} \sin\theta \cos\theta d\theta \int_0^{\omega_i^{Debye}} \hbar\omega \frac{dN_{i,j}(\omega, T)}{dT} d\omega. \quad (2.1)$$

where  $R$  is the TBR,  $c_{i,j}$  is the phonon propagation velocity in side  $i$  for phonons with mode  $j$  (longitudinal or transverse);  $\alpha_{i,j}$  is the transmission probability of phonons from side  $i$  with mode  $j$ ;  $\theta$  is the angle between the wave vector of the incident phonon and the normal to the interface; and  $\omega$  is the phonon frequency.  $N_{i,j}(\omega, T)$  is the density of phonons with energy on side  $i$  with mode  $j$  at temperature  $T$ . The Debye cutoff frequency is given by

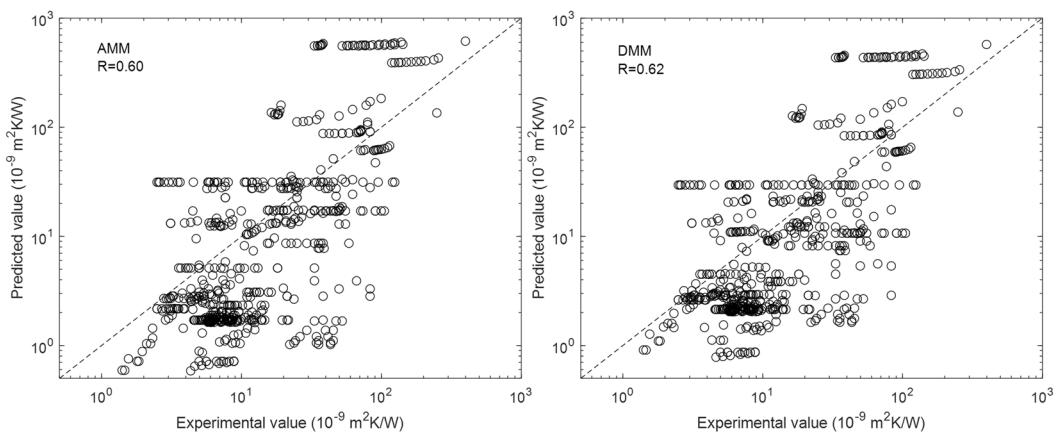
$$(\omega_i^{Debye})^3 = \frac{1}{\frac{1}{3} \sum_j \frac{1}{(c_{i,j})^3}} 6\pi^2 \left( \frac{N_{atom}}{V} \right)_i. \quad (2.2)$$

where  $N_{atom}/V$  is the atom number density of the crystal. For frequencies below  $\omega_i^{Debye}$

$$N_{i,j}(\omega, T) = \frac{\omega^2}{2\pi^2 c_{i,j}^3 [\exp(\hbar\omega/k_B T) - 1]}. \quad (2.3)$$

where  $k_B$  is the Boltzmann constant.

In the AMM,  $i=1$  in Eqs (2.1), (2.2), and (2.3), and the transmission probabilities of phonons are calculated from the acoustic impedances.



**Figure 1.** Correlation between the experimental values and the values predicted by the AMM and DMM.

$$\alpha_{1 \rightarrow 2} = \frac{4Z_2 Z_1}{(Z_1 + Z_2)^2} \quad (2.4)$$

where \$Z\_i = \rho\_i c\_i\$ is the acoustic impedance on each side of the interface, which equals to the product of the mass density and the phonon velocity.

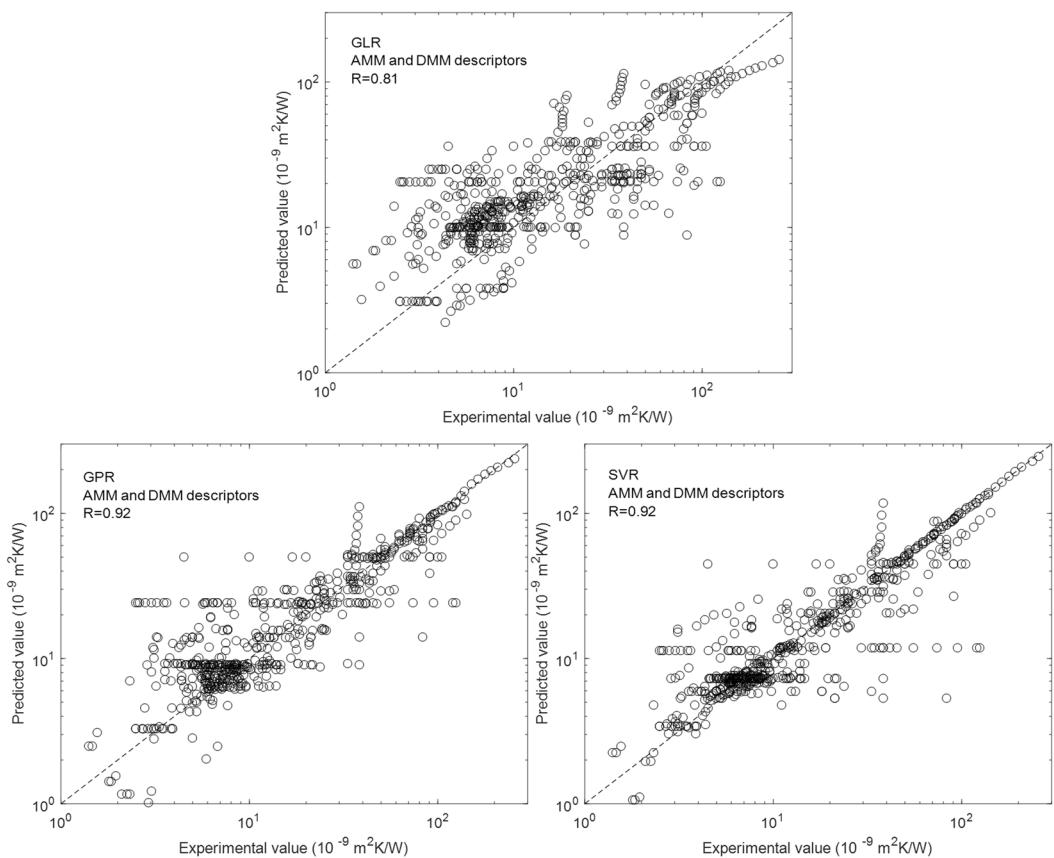
In the DMM, the transmission probability of phonons are calculated by

$$\alpha_{i,j}(\omega) = \frac{\sum_j c_{3-i,j}^{-2}}{\sum_{i,j} c_{i,j}^{-2}}. \quad (2.5)$$

The descriptors used for the AMM and DMM predictions are temperature, density, speed of sound (longitudinal and transverse), and unit cell volume, which we define as “AMM and DMM descriptors”. Figure 1 shows the correlation between the experimental values and the values predicted by the AMM and DMM. Two parameters, the correlation coefficient ( $R$ ) and root-mean-square error ( $RMSE$ ), are used to evaluate the predictive accuracy of the models. For the AMM,  $R$  and  $RMSE$  are 0.60 and 121.3, respectively. For the DMM,  $R$  and  $RMSE$  are 0.62 and 91.4, respectively. Thus both the AMM and DMM have low predictive accuracy, although the DMM is slightly better.

We next predict the TBR of all the interfaces using the machine learning models. It is known that the predictive accuracy of machine learning algorithms is strongly dependent on the selected descriptor sets<sup>77</sup>. In order to compare their predictions with those of the AMM and DMM, we have initially used the AMM and DMM descriptors. Figure 2 shows the correlation between the experimental and predicted values for the GLR, GPR, and SVR models. We next study the effects of descriptor set choice on the predictive accuracy of the machine learning models. First, we predict the TBR using all collected descriptors, because all collected measurement conditions and material properties may affect TBR. The values of  $R$  and  $RMSE$  for the GLR model and the average values of  $R$  and  $RMSE$  for the GPR and the SVM models are shown in Table 1. The results for the LASSO-GLR model are not given in Fig. 2 and Table 1, because the descriptors are different. As a result of model fitting, the LASSO-GLR model automatically selected 12 descriptors from the given 24, which are film thickness, film heat capacity, film melting point, film speed of sound (longitudinal), film thermal expansion, film unit cell volume, substrate thermal conductivity, substrate density, substrate speed of sound (longitudinal), substrate elastic modulus, substrate thermal expansion coefficient and substrate unit cell volume. We do not use this descriptor set in the following work because some important descriptors such as measurement temperature are missing, which is not reasonable from a physical point of view. The  $R$  and  $RMSE$  for the LASSO-GLR model are 0.89 and 15.8, respectively. For the GPR model, the average optimal  $\sigma_f$  and  $\sigma_l$  of the four random partitions are 176.22 and 5.56. For the SVR model, the average optimal  $C$ ,  $\epsilon$ , and  $\gamma$  are 987.15, 0.21 and 3.23. It is clear that all the machine learning models have much higher  $R$  and much lower  $RMSE$  than the AMM and DMM, indicating that all the machine learning models have a better predictive accuracy than the AMM and DMM. Among the machine learning models, the GPR and the SVR models have better performance than the GLR and the LASSO-GLR models.

However, model fitting using all collected descriptors is difficult because it is difficult to collect some of the descriptors, such as thermal conductivity, Debye temperature, speed of sound (longitudinal and transverse), elastic modulus, and bulk modulus, for all the materials from references. Also, different references usually give very different values for these material properties. For example, the measured thermal conductivity of crystalline AlN ceramics ranges from  $17 \text{ Wm}^{-1}\text{K}^{-1}$  to  $285 \text{ Wm}^{-1}\text{K}^{-1}$ <sup>78</sup>. The measured thermal conductivity of nanoscale AlN thin films could be as low as  $1.4 \text{ Wm}^{-1}\text{K}^{-1}$ <sup>79</sup>. Such large difference may significantly affect the predictive accuracy. Therefore, we predict TBR using only “reliable” descriptors: measurement temperature, film thickness, heat capacity, melting point, density, and unit cell volume. These material properties are called “reliable” because they are easily collected from references and different references usually give similar values. With the “reliable”



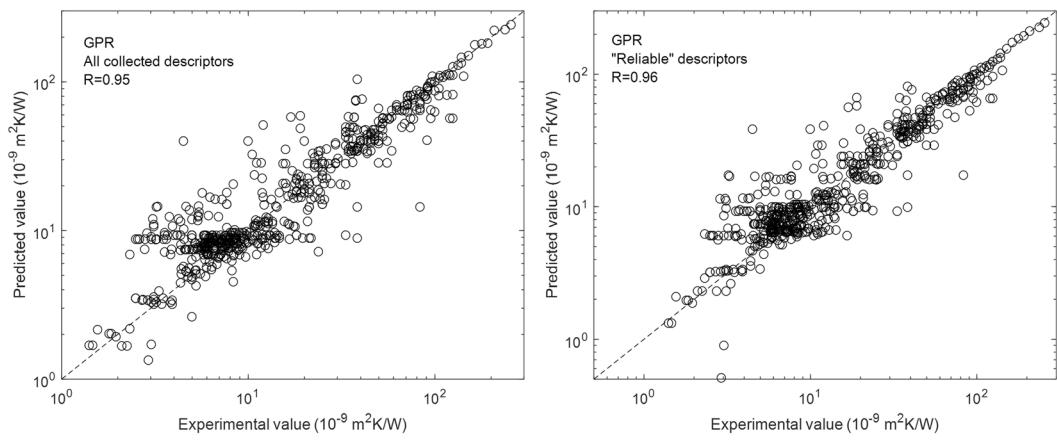
**Figure 2.** Correlation between the experimental values and the values predicted by the GLR, GPR, and SVR models using the AMM and DMM descriptors.

	AMM and DMM descriptors		All collected descriptors		“Reliable” descriptors	
	R	RMSE	R	RMSE	R	RMSE
AMM	0.60	121.3	N/A		N/A	
DMM	0.62	91.4	N/A		N/A	
GLR	0.81	20.3	0.9	15.4	0.81	20.1
GPR	0.92	13.2	0.96	9.6	0.96	9.4
SVR	0.92	13.9	0.95	10.4	0.96	9.9

**Table 1.** Comparison of R and RMSE predicted by different models using different sets of descriptors.

descriptors, we fit the GLR, GPR and the SVR models. The average optimal  $\sigma_f$  and  $\sigma_l$  are 106.68 and 2.59 for the GPR model. The average optimal  $C$ ,  $\varepsilon$ , and  $\gamma$  are 916.51, 0.88 and 1.71 for the SVR model. Like the use of all collected descriptors, it can also be found that the GPR and the SVR models have better predictive accuracies than the GLR model.

Figure 3 shows the correlation between the experimental values and the values predicted by GPR model using the “reliable” descriptors and all collected descriptors. It is clear that the predictions using the “reliable” descriptors and all collected descriptors have better predictive accuracy than using the AMM and DMM descriptors. This indicates that other descriptors, in addition to those that are used in the AMM and DMM, also affect TBR. If we compare the AMM and DMM descriptors and the “reliable” descriptors, we find that measurement temperature, density, and unit cell volume are in both descriptor sets. However, film thickness, heat capacity, and melting point are included in the “reliable” descriptors instead of speed of sound. This indicates film thickness, heat capacity, and melting point descriptors improved the predictive accuracy. In order to know which one plays the most important role, we added film thickness, heat capacity, or melting point to the AMM and DMM descriptor set and predicted TBR using the GPR model. The predictions had R values of 0.96, 0.92, and 0.92, respectively. This indicates that of these three descriptors, film thickness plays the most important role. To further test our suggestion, we predicted TBR using the GPR model and replacing film thickness in the “reliable” descriptor set with thermal conductivity, Debye temperature, elastic modulus, bulk modulus, or thermal expansion. All the



**Figure 3.** Correlation between the experimental values and the values predicted by the GPR model using the “reliable” descriptors and all collected descriptors.



**Figure 4.** Pearson correlation coefficient map between different materials properties. htcp (heat capacity), thcd (thermal conductivity), debye (Debye temperature), melt (melting point), dens (density), spdl (speed of sound longitudinal), spdt (speed of sound transverse), elam (elastic modulus), blkm (bulk modulus), thex (thermal expansion coefficient), and unitc (unit cell volume).

predictions without film thickness as descriptor show an  $R$  value of ~0.92, which is similar to that for the AMM and DMM descriptors and lower than the value of  $R$  when film thickness is included. These results indicate that film thickness is an important descriptor in the prediction of TBR. The dependence of TBR on film thickness has been found in our previous study of MD simulations on Si/Ge interface<sup>80</sup>. In that work, we used two different cell sizes, corresponding to film thickness of 50 and 100 nm, respectively. As a result, the 100 nm film has a lower TBR than the 50 nm one. The thickness dependence of TBR has also been observed in our experimental study on Au/sapphire interface<sup>26</sup>, where the TBR increases obviously when the grain size of Au decreases. The effect of film or grain size on TBR can be explained as that the phonon modes in the films are thickness or grain size dependent. Therefore, change in film thickness will change the phonon transmission probability at the interface as well as the TBR.

Furthermore, the predictions using the “reliable” descriptors and all collected descriptors have similar predictive accuracy. This indicates that the use of more material properties as descriptors does not necessarily improve predictive accuracy. It is known that many material properties, including heat capacity, thermal conductivity, Debye temperature, melting point, density, speed of sound, elastic modulus, bulk modulus, and thermal expansion coefficient, are physically-correlated. Using the collected materials property data, we drew a Pearson correlation coefficient map between different materials properties (see Fig. 4). For example, the correlation coefficients between Debye temperature and speed of sound are higher than 0.96, which is in good agreement with the physical-correlation of the two descriptors. We, therefore, attribute the similar predictive accuracy of the two descriptor sets to the physical-correlation of these descriptors; it is thus unnecessary to use all collected descriptors simultaneously in the prediction of TBR.

## Conclusions

In summary, we have demonstrated a machine learning method for predicting TBR. We found that machine learning models have better predictive accuracy than the commonly used AMM and DMM. Among the trained models, the GPR and the SVR models have better predictive accuracy. Also, by comparing the prediction results using different descriptor sets, we found that film thickness is an important descriptor in the prediction of TBR. These results indicate that machine learning is a simple, accurate, and cost-effective method for the prediction of TBR. It can, therefore, be used to guide the discovery of interfaces with very low or very high TBR for the thermal management of high power micro- and opto-electronic devices and for the development of high efficiency thermal barrier coatings and thermoelectric materials.

## References

- Pernot, G. *et al.* Precise control of thermal conductivity at the nanoscale through individual phonon-scattering barriers. *Nat. Mater.* **9**, 491–495 (2010).
- O'Brien, P. J. *et al.* Bonding-induced thermal conductance enhancement at inorganic heterointerfaces using nanomolecular monolayers. *Nat. Mater.* **12**, 118–122 (2013).
- Loegel, M. D., Grady, M. E., Sottos, N. R., Cahill, D. G. & Braun, P. V. Effects of chemical bonding on heat transport across interfaces. *Nat. Mater.* **11**, 502–506 (2012).
- Holland, M. G. Phonon scattering in semiconductors from thermal conductivity studies. *Phys. Rev.* **134**, A471–A480 (1964).
- Regner, K. T. Broadband phonon mean free path contributions to thermal conductivity measured using frequency domain thermoreflectance. *Nat. Commun.* **1640**, 1–7 (2013).
- Swartz, E. T. & Pohl, R. O. Thermal boundary resistance. *Rev. Mod. Phys.* **65**, 605–668 (1989).
- Landry, E. S. & McGaughey, A. J. H. Thermal boundary resistance predictions from molecular dynamics simulations and theoretical calculations. *Phys. Rev. B* **80**, 165304 (2009).
- Stevens, R. J., Zhigilei, L. V. & Norris, P. M. Effects of temperature and disorder on thermal boundary conductance at solid–solid interfaces: Nonequilibrium molecular dynamics simulations. *Int. J. Heat Mass Transfer* **50**, 3977–3989 (2007).
- Meller, J. *Encyclopedia of Life Sciences*, Nature Publishing Group (2001).
- Langley, P. & Simon, H. A. Applications of machine learning and rule induction. *Commun ACM* **38**, 1 (1995).
- Hopkins, P. E., Duda, J. C., Petz, C. W. & Floro, J. A. Controlling thermal conductance through quantum dot roughening at interfaces. *Phys. Rev. B* **84**, 035438 (2011).
- Hopkins, P. E., Phinney, L. M., Serrano, J. R. & Beechem, T. E. Effects of surface roughness and oxide layer on the thermal boundary conductance at aluminum/silicon interfaces. *Phys. Rev. B* **82**, 085307 (2010).
- Costescu, R. M., Wall, M. A. & Cahill, D. G. Thermal conductance of epitaxial interfaces. *Phys. Rev. B* **67**, 054302 (2003).
- Lyeo, H.-K. & Cahill, D. G. Thermal conductance of interfaces between highly dissimilar materials. *Phys. Rev. B* **73**, 144301 (2006).
- Sakata, M. *et al.* Thermal conductance of silicon interfaces directly bonded by room-temperature surface activation. *Appl. Phys. Lett.* **106**, 081603 (2015).
- Stevens, R. J., Smith, A. N. & Norris, P. M. Measurement of thermal boundary conductance of a series of metal–dielectric interfaces by the transient thermoreflectance technique. *J. Heat Transfer* **127**(3), 315–322 (2005).
- Koh, Y. K., Bae, M.-H., Cahill, D. G. & Pop, E. Heat conduction across monolayer and few-layer graphenes. *Nano Lett.* **10**(11), 4363–4368 (2010).
- Chen, Z., Jang, W., Bao, W., Lau, C. N. & Dames, C. Thermal contact resistance between graphene and silicon dioxide. *Appl. Phys. Lett.* **95**, 161910 (2009).
- Hopkins, P. E. & Norris, P. M. Thermal boundary conductance response to a change in Cr/Si interfacial properties. *Appl. Phys. Lett.* **89**, 131909 (2006).
- Hopkins, P. E. *et al.* Effect of dislocation density on thermal boundary conductance across GaSb/GaAs interfaces. *Appl. Phys. Lett.* **98**, 161913 (2011).
- Collins, K. C., Chen, S. & Chen, G. Effects of surface chemistry on thermal conductance at aluminum–diamond interfaces. *Appl. Phys. Lett.* **97**, 083102 (2010).
- Stoner, R. J. & Maris, H. J. Kapitza conductance and heat flow between solids at temperatures from 50 to 300 K. *Phys. Rev. B* **48**, 16373 (1993).
- Wang, R. Y., Segalman, R. A. & Majumdar, A. Room temperature thermal conductance of alkanedithiol self-assembled monolayers. *Appl. Phys. Lett.* **89**, 173113 (2006).
- Ge, Z., Cahill, D. G. & Braun, P. V. Thermal Conductance of Hydrophilic and Hydrophobic Interfaces. *Phys. Rev. Lett.* **96**, 186101 (2006).
- Shenogina, N., Godawat, R., Kebblinski, P. & Garde, S. How wetting and adhesion affect thermal conductance of a range of hydrophobic to hydrophilic aqueous interfaces. *Phys. Rev. Lett.* **102**, 156101 (2009).
- Xu, Y., Kato, R. & Goto, M. Effect of microstructure on Au/sapphire interfacial thermal resistance. *J. Appl. Phys.* **108**, 104317 (2010).
- Xu, Y., Wang, H., Tanaka, Y., Shimono, M. & Yamazaki, M. Measurement of interfacial thermal resistance by periodic heating and a thermo-reflectance technique. *Mater. Trans.* **48**, 148–150 (2007).
- Zhan, T. *et al.* Thermal boundary resistance at Au/Ge/Ge and Au/Si/Ge interfaces. *RSC Adv.* **5**, 49703–49707 (2015).
- Zhan, T. *et al.* Phonons with long mean free paths in a-Si and a-Ge. *Appl. Phys. Lett.* **104**, 071911 (2014).
- Xu, Y., Goto, M., Kato, R., Tanaka, Y. & Kagawa, Y. Thermal conductivity of ZnO thin film produced by reactive sputtering. *J. Appl. Phys.* **111**, 084320 (2012).
- Kato, R., Xu, Y. & Goto, M. Development of a frequency-domain method using completely optical techniques for measuring the interfacial thermal resistance between the metal film and the substrate. *Jpn. J. Appl. Phys.* **50**, 106602 (2011).
- Kato, R. & Hatta, I. Thermal conductivity and interfacial thermal resistance: measurements of thermally oxidized SiO<sub>2</sub> films on a silicon wafer using a thermo-reflectance technique. *Int. J. Thermophys.* **29**, 2062–2071 (2008).
- Hopkins, P. E. *et al.* Manipulating thermal conductance at metal–graphene contacts via chemical functionalization. *Nano Lett.* **12**, 590–595 (2012).
- Hsieh, W. P., Lyons, A. S., Pop, E., Kebblinski, P. & Cahill, D. G. Pressure tuning of the thermal conductance of weak interfaces. *Phys. Rev. B* **84**, 184107 (2011).
- Cai, W. *et al.* Thermal transport in suspended and supported monolayer graphene grown by chemical vapor deposition. *Nano Lett.* **10**, 1645–1651 (2010).
- Schmidt, A. J., Collins, K. C., Minnich, A. J. & Chen, G. Thermal conductance and phonon transmissivity of metal-graphite interfaces. *J. Appl. Phys.* **107**, 104907 (2010).
- Huxtable, S. T. *et al.* Interfacial heat flow in carbon nanotube suspensions. *Nat. Mater.* **2**, 731–734 (2003).
- Kim, E.-K., Kwun, S.-I., Lee, S.-M., Seo, H. & Yoon, J.-G. Phonon and electron transport through Ge<sub>2</sub>Sb<sub>2</sub>Te<sub>5</sub> films and interfaces bounded by metals. *Appl. Phys. Lett.* **76**, 3864 (2000).
- Hopkins, P. E., Norris, P. M. & Stevens, R. J. Influence of inelastic scattering at metal–dielectric interfaces. *J. Heat Transfer* **130**(2), 022401 (2008).

40. Hopkins, P. E., Norris, P. M., Stevens, R. J., Beechem, T. E. & Graham, S. Influence of interfacial mixing on thermal boundary conductance across a chromium/silicon interface. *J. Heat Transfer* **130**(6), 062402 (2008).
41. Zheng, H. & Jagannadham, K. Thermal conductivity and interface thermal conductance in composites of titanium with graphene platelet. *J. Heat Transfer* **136**(6), 061301 (2014).
42. Hu, C., Kiene, M. & Ho, P. S. Thermal conductivity and interfacial thermal resistance of polymeric low  $k$  films. *Appl. Phys. Lett.* **79**, 4121 (2001).
43. Minnich, A. J. *et al.* Thermal conductivity spectroscopy technique to measure phonon mean free paths. *Phys. Rev. Lett.* **107**, 095901 (2011).
44. Hopkins, P. E. *et al.* Reduction in thermal boundary conductance due to proton implantation in silicon and sapphire. *Appl. Phys. Lett.* **98**, 231901 (2011).
45. Monachon, C., Hojeij, M. & Weber, L. Influence of sample processing parameters on thermal boundary conductance value in an Al/AlN system. *Appl. Phys. Lett.* **98**, 091905 (2011).
46. Gorham, C. S. *et al.* Ion irradiation of the native oxide/silicon surface increases the thermal boundary conductance across aluminum/silicon interfaces. *Phys. Rev. B* **90**, 024301 (2014).
47. Gengler, J. J. *et al.* Limited thermal conductance of metal–carbon interfaces. *J. Appl. Phys.* **112**, 094904 (2012).
48. Freedman, J. P., Yu, X., Davis, R. F., Gellman, A. J. & Malen, J. A. Thermal interface conductance across metal alloy–dielectric interfaces. *Phys. Rev. B* **93**, 035309 (2016).
49. Jeong, M. *et al.* Enhancement of thermal conductance at metal–dielectric interfaces using subnanometer metal adhesion layers. *Phys. Rev. Appl.* **5**, 014009 (2016).
50. Donovan, B. F. *et al.* Thermal boundary conductance across metal–gallium nitride interfaces from 80 to 450 K. *Appl. Phys. Lett.* **105**, 203502 (2014).
51. Cho, J. *et al.* Phonon scattering in strained transition layers for GaN heteroepitaxy. *Phys. Rev. B* **89**, 115301 (2014).
52. Gaskins, J. T. *et al.* Thermal conductance across phosphonic acid molecules and interfaces: ballistic versus diffusive vibrational transport in molecular monolayers. *J. Phys. Chem. C* **119**(36), 20931 (2015).
53. Hopkins, P. E. *et al.* Influence of anisotropy on thermal boundary conductance at solid interfaces. *Phys. Rev. B* **84**, 125408 (2011).
54. Chow, P. K. *et al.* Gold-titania interface toughening and thermal conductance enhancement using an organophosphonate nanolayer. *Appl. Phys. Lett.* **102**, 201605 (2013).
55. Norris, P. M., Smoyer, J. L., Duda, J. C. & Hopkins, P. E. Prediction and measurement of thermal transport across interfaces between isotropic solids and graphitic materials. *J. Heat Transfer* **134**(2), 020910 (2011).
56. Koh, Y. K., Cao, Y., Cahill, D. G. & Jena, D. Heat-transport mechanisms in superlattices. *Adv. Funct. Mater.* **19**, 610–615 (2009).
57. Costescu, R. M., Cahill, D. G., Fabreguette, F. H., Sechrist, Z. A. & George, S. M. Ultra-low thermal conductivity in W/Al<sub>2</sub>O<sub>3</sub> nanolaminates. *Science* **303**, 989 (2004).
58. Cheaito, R. *et al.* Thermal boundary conductance accumulation and interfacial phonon transmission: Measurements and theory. *Phys. Rev. B* **91**, 035432 (2015).
59. Shukla, N. C. *et al.* Thermal conductivity and interface thermal conductance of amorphous and crystalline Zr<sub>47</sub>Cu<sub>31</sub>Al<sub>13</sub>Ni<sub>9</sub> alloys with a Y<sub>2</sub>O<sub>3</sub> coating. *Appl. Phys. Lett.* **94**, 081912 (2009).
60. Ma, Y. Hotspot size-dependent thermal boundary conductance in nondiffusive heat conduction. *J. Heat Transfer* **137**(8), 082401 (2015).
61. Losego, M. D., Moh, L., Arpin, K. A., Cahill, D. G. & Braun, P. V. Interfacial thermal conductance in spun-cast polymer films and polymer brushes. *Appl. Phys. Lett.* **97**, 011908 (2010).
62. Putnam, S. A., Cahill, D. G., Ash, B. J. & Schadler, L. S. High-precision thermal conductivity measurements as a probe of polymer/nanoparticle interfaces. *J. Appl. Phys.* **94**, 6785 (2003).
63. Jin, Y., Shao, C., Kieffer, J., Pipe, K. P. & Shteiin, M. Origins of thermal boundary conductance of interfaces involving organic semiconductors. *J. Appl. Phys.* **112**, 093503 (2012).
64. Lee, S.-M. & Cahill, D. G. Heat transport in thin dielectric films. *J. Appl. Phys.* **81**, 2590 (1997).
65. Szwejkowski, C. J. *et al.* Size effects in the thermal conductivity of gallium oxide ( $\beta$ -Ga<sub>2</sub>O<sub>3</sub>) films grown via open-atmosphere annealing of gallium nitride. *J. Appl. Phys.* **117**, 084308 (2015).
66. Clark, S. P. R. *et al.* Growth and thermal conductivity analysis of polycrystalline GaAs on chemical vapor deposition diamond for use in thermal management of high-power semiconductor lasers. *J. Vac. Sci. Technol. B* **29**, 03C130 (2011).
67. Foley, B. M. *et al.* Modifying surface energy of graphene via plasma-based chemical functionalization to tune thermal and electrical transport at metal interfaces. *Nano Lett.* **15**(8), 4876–4882 (2015).
68. Duda, J. C. *et al.* Influence of interfacial properties on thermal transport at gold:silicon contacts. *Appl. Phys. Lett.* **102**, 081902 (2013).
69. Duda, J. C. & Hopkins, P. E. Systematically controlling Kapitza conductance via chemical etching. *Appl. Phys. Lett.* **100**, 111602 (2012).
70. Inorganic Material Database (AtomWork), NIMS Materials Database (MatNavi).
71. Touloukian, Y. S. *Thermophysical properties of matter, the TPRC data series*, Plenum Publishing Corporation (1970).
72. Dobson, A. J. & Barnett, A. *An introduction to generalized linear models*, Third Edition, Chapman and Hall/CRC (2008).
73. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B* **58**, 267–288 (1996).
74. Rasmussen, C. E. & Williams, C. K. I. *Gaussian processes for machine learning*. MIT Press. Cambridge, Massachusetts (2006).
75. Lu, W. C. *et al.* Using support vector machine for materials design. *Adv. Manuf.* **1**, 151–159 (2013).
76. The Mathworks, Inc. Statistics and machine learning toolbox, MATLAB, R2016a.
77. Hansen, K. *et al.* Machine learning predictions of molecular properties: Accurate many-body potentials and non-locality in chemical space. *J. Phys. Chem. Lett.* **6**, 2326 (2015).
78. Júnior, A. F. Thermal conductivity of polycrystalline aluminum nitride (AlN) ceramics. *Cerâmica* **50**, 247–253 (2004).
79. Zhao, Y. *et al.* Pulsed photothermal reflectance measurement of the thermal conductivity of sputtered aluminum nitride thin films. *J. Appl. Phys.* **96**, 4563–4568 (2004).
80. Zhan, T., Minamoto, S., Xu, Y., Tanaka, Y. & Kagawa, Y. Thermal boundary resistance at Si/Ge interfaces by molecular dynamics simulation. *AIP Advances* **5**, 047102 (2015).

## Acknowledgements

This work was supported by the “Materials research by Information Integration” Initiative (MI<sup>2</sup>I) project of the Support Program for Starting Up Innovation Hub from Japan Science and Technology Agency (JST).

## Author Contributions

Y.X. designed and supervised the study. T.Z. conducted the data collection. L.F. conducted the analysis. T.Z. and L.F. wrote the manuscript text.

## Additional Information

**Supplementary information** accompanies this paper at doi:[10.1038/s41598-017-07150-7](https://doi.org/10.1038/s41598-017-07150-7)

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017