To Reviewer # 6:

We are appreciative of the constructive feedback from the reviewer. As suggested by the reviewer, we have extended the current example.

To Comment # 1:

Concern: ***Fig. 3 may be deleted.***

Response: Yes, we agree with the reviewer. We have, accordingly, deleted Fig. 3 and removed some contents in Section 3 that are available in the literature.

To Comment # 2:

Concern: ***Check the expression of "due to the spatial variability of the inputs" at the right column of Page 3.***

Response: Yes, we have checked the expression and made appropriate changes.

To Comment # 3:

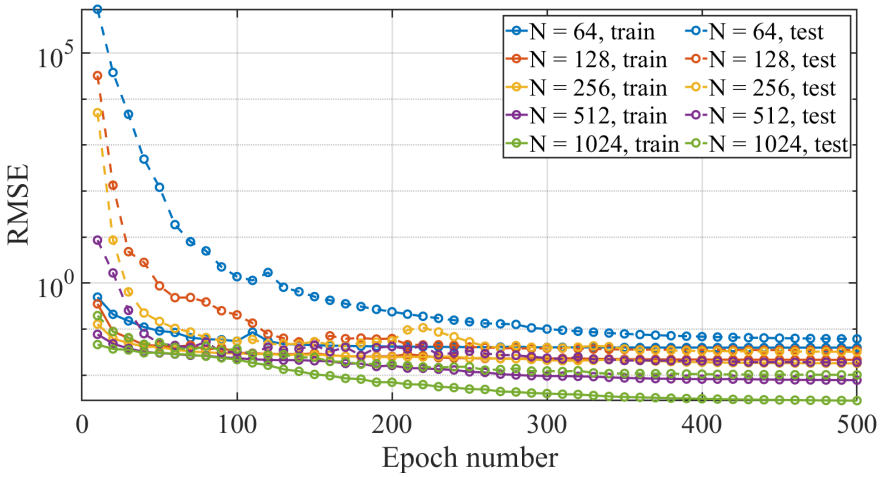
Concern: ***Check the expression of "It is assumed that each i,i = 1; 2; l1 extracts k feature maps. There are k and K in the manuscript."***

Response: Yes, thank you for pointing this out. We have checked the use of uppercase and lowercase letters. In the revised version, k denotes the growth rate of each block and there is no K.

To Comment # 4:

Concern: ***What is unit and title of the vertical axis of Fig. 6(a)?***

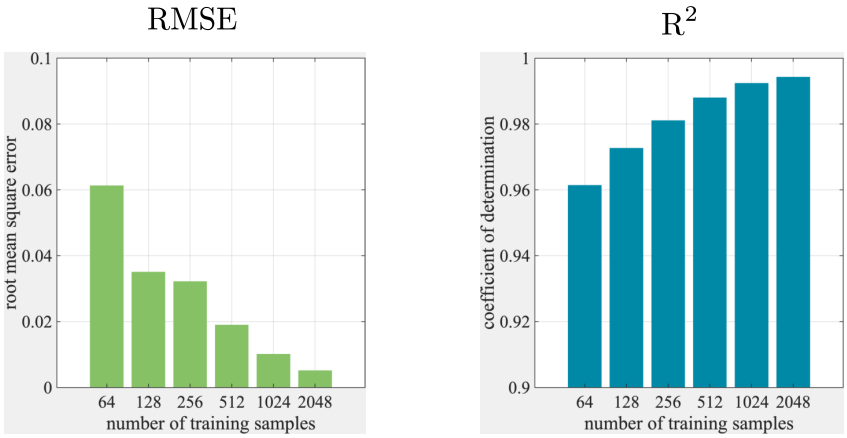
Response: Yes, we have added the axis label to Fig. 6(a), Fig. 9(a), and Fig. 12(a), respectively. It shows the convergence of the test and training RMSE where we vary the number of samples. The computed RMSE value by definition has the same units as the quantity being estimated. In the data preprocessing phase, we standardized the inputs and outputs. Hence, the calculated RMSE value is a dimensionless variable. But it serves as an informative indicator in light of the optimization process, that is, a lower value of RMSE indicates better model fit.



To Comment # 5:

Concern: ***Is it possible to reduce values of RMSE by increasing the training size of 2048 or more?***

Response: Yes, the RMSE value will decrease when 2048 samples are used. We have also checked the R2 value. Note the improvement of prediction performance is subtle. This is due to the bias and variance associated with the predictive model. Because the bias and variance are the two components of imprecision, there is a trade-off between them in general. To address the bias issue, i.e. underfitting, we added sufficient layers to the deep learning model, ensuring the prediction performance on the training and testing data. To address the variance issue, i.e. overfitting, we adopted techniques such as dropout, early stopping, and cross-validation in addtion to the L2 regularization.



To Comment # 6:

Concern: ***Verification of the proposed approach using a more complex example is necessary.***

Response: Yes, we understand the importance of model generalizability and capability regarding surrogate modeling. To demonstatre the generalizability, uncertainty quantification of Poisson's

equation and Darcy's law has been carried out. To demonstatre the advanced capability, geometric nonlinearity has been integrated into the current case study. Enclosed please find the results file (file name: summary.pdf). Morever, a detailed discussion on potential applications have been added to Section 5 of the revised manuscript. To support more applications, computer codes will be made available at `https://xihaier.github.io` upon publication of this manuscript.