Machine learning modeling of power delivery networks with varying decoupling capacitors

Yeong Kang Liew¹, Nur Syazreen Ahmad², Azniza Abd Aziz², Patrick Goh²

¹Intel Corporation, Penang, Malaysia

²School of Electrical and Electronic Engineering, Universiti Sains Malaysia, Nibong Tebal, Malaysia

Article Info	ABSTRACT		
Article history:	This paper presents modeling of power delivery network (PDN) impedance with vary-		
Received Jul 8, 2021 Revised Mar 8, 2022 Accepted Apr 5, 2022	ing decoupling capacitor placements using machine learning techniques. The use of multilayer perceptron artificial neural networks (ANN) and gaussian process regression (GPR) techniques are explored, and the effects of the hyperparameters such as the number of hidden neurons in the ANN, and the choice of kernel functions in the GPR are investigated. The best performing networks in each case are selected and		
Keywords:	compared in terms of accuracy using test data consisting of PDN impedance responses		
Artificial neural network Gaussian process regression Power delivery network	that were never encountered during training. Results show that the GPR models were significantly more accurate than the ANN models, with an average mean absolute error of 5.23 m Ω compared to 11.33 m Ω for the ANN.		

This is an open access article under the <u>CC BY-SA</u> license.



Corresponding Author:

Patrick Goh School of Electrical and Electronic Engineering, Universiti Sains Malaysia 14300, Nibong Tebal, Penang, Malaysia Email: eepatrick@usm.my

1. INTRODUCTION

Power integrity (PI) is a field of engineering which strives to ensure the quality of power delivered to integrated circuit components through a network known as the power delivery network (PDN) [1], [2]. Poorly designed PDN can induce serious noise problems to the board and affect the signal quality, which can impair data throughput and results in reduced performances. Good PDN designs should satisfy the given target impedances in a circuit in order to minimize direct current (DC) effects such as IR drop, and also alternating current (AC) effects such as simultaneous switching noise and ground bounce [3].

To minimize the impedance of a PDN, one of the most common approach is to place decoupling capacitors on the PDN, especially at locations close to the power pins of the devices on the network [4], [5]. The decoupling capacitors are used to provide low impedance paths from the power source to the power sinks and also to filter or guide the noise away to the ground. However, the design and placements of these decoupling capacitors on a PDN is not a trivial task. Too few or poorly placed decoupling capacitors would result in PDNs which fail to meet the target specifications, while too many decoupling capacitors would increase the size and cost of the design. In addition, all physical capacitors have associated parasitic inductances, which are negligeable at lower frequencies, but can dominate the performance at higher frequencies, depending on the value of the capacitors. Thus, proper sizing of the capacitors also play an important role in satisfying the target impedance over the range of the operating frequency in the design.

Many previous work exists on optimizing the placement, sizing, and selection of decoupling capacitors on a PDN [6]-[9]. However, these work are based on theories and rule of thumbs which can vary from design

to design and are usually not easy to generalize and implement practically, without an expert knowledge in the field. In addition, they require extensive time and effort to study and optimize the design repeatedly, often relying on computationally expensive computer-aided design simulations [10]-[12].

Recently, machine learning (ML) techniques have seen considerable success in various engineering fields [13], [14], particularly also in the closely related signal integrity (SI) field [15]-[18]. As a result, researchers have explored the use of ML techniques for the solution of PI and PDN problems [19]-[21]. This includes work utilizing particle swarm optimization, genetic algorithm, and Q-learning to optimize the placement, sizing, and selection of the decoupling capacitors on a PDN [22]-[26]. The general purpose of these research works is to optimize the design to lower down the input impedance such that it is lower than the target or desired impedance value with the least number of decoupling capacitors used in PDN. However, while these researches focus on the optimization of the PDN, they do not help in the impedance response prediction of these designs. This is important as well as designers often prefer to look at the actual impedance response plots of the PDNs to gain insight into their performances. The work in [27] applies a recurrent neural network to predict the impedance response curve but it is related to extrapolations of the impedance response beyond the training frequency by learning the correlation of the existing impedance values. Thus, this method cannot be applied directly to obtain the impedance response from a physical design.

This paper presents a study of ML techniques to model and predict the impedance response curves of PDNs with varying decoupling capacitor placements. Artificial neural network (ANN) and Gaussian process regression (GPR) are applied to model the PDN from a physical design with different decoupling capacitor model and values. Effects of different training methods in the ANN and kernel functions in the GPR are also investigated, and the accuracy of the various networks in both methods are compared. The fundamentals of ANN and GPR are assumed to be well known to the reader, and are not repeated in this paper for the benefit of brevity.

2. RESEARCH METHOD

2.1. Data generation

The data used in this work is obtained by simulation of actual impedance profiles of power delivery networks obtained from Intel Corporation. The PDN is driven by a voltage regulator module (VRM) and has a total of 17 locations for decoupling capacitor placements. Four different decoupling capacitor models are considered and this is tabulated in Table 1. Due to the different models and location proximity to the VRM on the PDN, each decoupling capacitor location is restricted to only receive one type of capacitor. However, note that it is not necessary for a location to have a capacitor, and instead, it can be left empty if necessary. Table 2 tabulates the possible locations for each of the decoupling capacitor, where the 17 locations are labeled from A to Q. Next, the impedance profiles of the PDN with varying decoupling capacitor placements are simulated in Cadence Allegro Sigrity OptimizePI.

ID	Model name	Manufacturer	Model type	Capacitance (μF)	Area (mil ²)	Rated	Self resonance
						voltage (V)	freq. (MHz)
1	GRM155R61E104KA87	Murata	SPICE	0.1	800	25	26.0
2	LMK105BBJ475MVLF	Taiyo Yuden	SPICE	4.7	800	10	3.98
3	CL05A106MQ5NUN	Samsung	SPICE	10	800	6.3	2.51
4	GRM21BR60J476ME15	Murata	SPICE	47	4000	6.3	1.00

Table 1. Decoupling capacitor models used in this work

Table 2. Possible locations for each of the capacitor model

Capacitor ID	Possible locations	
1	D, F, H, I, J, K, L, M, N, O	
2	E, G	
3	А	
4	B, C, P, Q	

Two different sampling methods are considered in this work. In the first method, a fine sampling is performed with a total of 236 points over the frequency range of 10 kHz to 200 MHz, while in the second method, a coarse sampling is performed with a total of 29 points over the frequency range of 100 kHz to

200 MHz. The fine sampling is naturally more accurate, but is also computationally more expensive in both the training data generation and machine learning training processes. A comparison will be drawn based on the results in both sampling methods in the next section. Once the training data has been generated based on the two sampling methods, two machine learning modeling techniques are investigated to model the impedance profiles of the PDN. The first utilizes a multilayer perceptron neural network, while the second utilizes a Gaussian process regression model. The following subsections describe the process and parameter used in both methods.

2.2. ANN modeling method

A three layered multilayer perceptron neural network is used model the PDN to predict the impedance response over the frequency range of interest. The inputs are the decoupling capacitor placements and combinations, while the outputs are the impedance response of the PDN. A hyperbolic tangent sigmoid function is used as the activation function for the input layer, while a linear activation function is used for the output layer. To avoid overweighting the larger data values, the input and output data values are normalized to be between -1 and 1. Two different backpropagation training algorithms are considered, which are the Levenberg-Marquardt and the Bayesian regularization training functions. Early stopping is used to prevent overfitting. In order to determine the optimal number of hidden neurons, a linear sweep is performed from 10-50 hidden neurons, with an increment of 5 neurons per iteration. Finally, the best performing neural network is selected based on the lowest mean absolute error.

2.3. GPR modeling method

The GPR model has the same modeling goals, and input and output targets as the ANN in the previous subsection. The main hyperparameter in the GPR is the choice of kernel or covariance function. For that, 10 different kernel functions are considered, ranging from exponential, squared exponential, Matérn with parameters 3/2 and 5/2, and rational quadratic kernels, and all are considered with the same or a separate length scale per predictor. Thus, 10 separate GPR models are trained, one for each kernel. In each case, Bayesian optimization is used to optimize the remaining hyperparameters by exploring the search space and minimizing the cross-validation loss. Finally, the best performing GPR model is selected based on the lowest mean absolute error.

3. RESULT AND DISCUSSION

3.1. ANN modeling results

The mean absolute error (MAE) for the ANN trained using the Levenberg-Marquardt and Bayesian regularization training algorithms with varying number of hidden neurons are shown in Table 3 for the case of a fine sampling with 236 points over the frequency range of interest, and in Table 4 for the case of a coarse sampling with 29 points over the frequency range of interest. It can be seen that the fine sampling produces a more consistent result across all the number of hidden neurons, while the results from the coarse sampling shows more fluctuations. This is due to the fact that the coarse sampling data has less overall points across the frequency range, and thus more variability between the points. However, the coarse sampling data models can yield better results than the fine sampling data models at the optimal number of hidden neurons. The best performing network is selected as the ANN with 35 hidden neurons trained with Bayesian regularization for the fine sampling model, and the ANN with 20 hidden neurons trained with Levenberg-Marquardt for the coarse sampling model.

No. of hidden neurons	Levenberg-marquardt MAE (m Ω)	Bayesian regularization MAE (m Ω)
10	15.80	14.83
15	15.97	14.93
20	15.30	14.60
25	15.37	14.80
30	14.93	14.93
35	15.27	14.43
40	15.97	14.77
45	14.97	14.90
50	14.77	14.80

Table 3. MAE for the ANN models with varying number of hidden neurons for fine sampling

No. of hidden neurons	Levenberg-marquardt MAE (m Ω)	Bayesian regularization MAE (m Ω)
10	35.2	11.4
15	9.1	11.0
20	5.6	8.6
25	22.6	9.2
30	16.6	30.3
35	31.1	9.5
40	41.9	87.1
45	10.1	34.7
50	21.4	24.2

Table 4. MAE for the ANN models with varying number of hidden neurons for coarse sampling

3.2. GPR modeling results

The MAE for the GPR trained using different kernel functions are shown in Table 5 for the case of the fine sampling and Table 6 for the case of the coarse sampling. It can be seen that the GPR models show less fluctuations with varying kernel functions, compared to the ANNs with varying number of hidden neurons. In addition, using a same length scale per predictor produces better results on average than using a separate length scale per predictor. The best performing model is selected as the GPR with an exponential kernel function and a same length scale for each predictor, in both cases.

Table 5. MAE for the GPR models with varying kernel functions for fine sampling

Kernel function	Length scale per predictor	MAE $(m\Omega)$
Squared exponential	Same	15.27
Exponential	Same	12.07
Matern 3/2	Same	15.63
Matern 5/2	Same	15.60
Rational quadratic	Same	15.50
Squared exponential	Separate	58.00
Exponential	Separate	58.00
Matern 3/2	Separate	58.00
Matern 5/2	Separate	58.00
Rational quadratic	Separate	58.00

Table 6. MAE for the GPR models with varying kernel functions for coarse sampling

Kernel function	Length scale per predictor	MAE $(m\Omega)$
Squared exponential	Same	5.07
Exponential	Same	4.80
Matern 3/2	Same	5.00
Matern 5/2	Same	5.03
Rational quadratic	Same	5.00
Squared exponential	Separate	24.37
Exponential	Separate	24.47
Matern 3/2	Separate	21.50
Matern 5/2	Separate	24.57
Rational quadratic	Separate	23.80

3.3. Comparison between ANN and GPR

In order to perform an unbiased testing and comparison between the ANN and GPR models, two additional PDN impedance profiles are simulated using Cadence Allegro Sigrity OptimizePI with decoupling capacitor placement combinations that were never seen during training. The best performing models of the ANN and GPR depicted in the previous subsections were selected, and the models were used to predict the impedance profiles in both the fine and coarse sampling frequency range. Figure 1 shows a comparison of the predicted results from the ANN and GPR models for test case 1, where Figure 1(a) shows the result using fine sampling, while Figure 1(b) shows the result using coarse sampling. Similarly, Figure 2 shows a comparison of the predicted results from the ANN and GPR models for test case 2, where Figure 2(a) shows the result using fine sampling, while Figure 2(b) shows the result using coarse sampling.

From the figures, it can be seen that the GPR models produce the more accurate prediction on the testing datasets, for both the fine and coarse sampling data. While the performance of both the ANN and

GPR models during training were comparable to each other, the GPR is able to generalize much better when presented with new datasets that were not encountered during training. This can be attributed to the non-parametric nature of the GPR that predicts outcomes in a probabilistic manner, compared to the parametric ANNs. The MAE values for both models on the testing datasets are summarized in Table 7.



Figure 1. Comparison of the impedance profile predicted results from the ANN and GPR models for test case 1 (a) fine sampling and (b) coarse sampling



Figure 2. Comparison of the impedance profile predicted results from the ANN and GPR models for test case 2 (a) fine sampling and (b) coarse sampling

Table 7. MAE for the ANN and GPR models for the testing data					
Model	Test case 1 MAE (m Ω)		Test case 2 MAE (m Ω)		
	Fine sampling	Coarse sampling	Fine sampling	Coarse sampling	
ANN	14.1	12.4	5.8	13.0	
GPR	8.3	2.5	6.2	3.9	

4. CONCLUSION

In this paper, machine learning methods are explored for the modeling of power delivery network impedance with varying decoupling capacitor placements. The applications of artificial neural networks and

Gaussian process regression techniques were demonstrated, and the effects of the number of hidden neurons in ANN, and the kernel function in GPR were investigated. The best performing networks were compared in terms of the accuracy in modeling a real PDN for both a fine and a coarse sampling frequency range. It is found that the GPR greatly outperforms the ANN in this regard with an average MAE of 5.23 m Ω compared to 11.33 m Ω for the ANN, on new test data that were never encountered during training. Future work will focus on optimization of search space exploration techniques to improve the training data generation process, such that it can better represent the PDN and frequency range of interest. In addition, improved hyperparameter optimization techniques can also be pursued to further improve the generalization ability of the models.

ACKNOWLEDGEMENT

This work is supported by the Ministry of Higher Education, Malaysia, under the Fundamental Research Grant Scheme (FRGS) grant number FRGS/1/2020/TK0/USM/02/7.

REFERENCES

- [1] J. Fan, "Signal integrity and power integrity," *IEEE Electromagnetic Compatibility Magazine*, vol. 9, no. 2, pp. 60–60, 2020, doi: 10.1109/memc.2020.9133245.
- [2] Z. Yang, "Fundamentals of power integrity," in 2018 IEEE Symposium on Electromagnetic Compatibility, Signal Integrity and Power Integrity (EMC, SI and PI), 2018, pp. 1–50, doi: 10.1109/EMCSI.2018.8495347.
- [3] M. Swaminathan, D. Chung, S. Grivet-Talocia, K. Bharath, V. Laddha, and J. Xie, "Designing and modeling for power integrity," *IEEE Transactions on Electromagnetic Compatibility*, vol. 52, no. 2, pp. 288–310, May 2010, doi: 10.1109/TEMC.2010.2045382.
- [4] J. N. Tripathi, J. Mukherjee, P. R. Apte, N. K. Chhabra, R. K. Nagpal, and R. Malik, "Selection and placement of decoupling capacitors in high speed systems," *IEEE Electromagnetic Compatibility Magazine*, vol. 2, no. 4, pp. 72–78, 2013, doi: 10.1109/MEMC.2013.6714703.
- [5] T. Wu, H. Chuang, and T. Wang, "Overview of power integrity solutions on package and PCB: decoupling and EBG isolation," *IEEE Transactions on Electromagnetic Compatibility*, vol. 52, no. 2, pp. 346–356, May 2010, doi: 10.1109/TEMC.2009.2039575.
- [6] K. Koo, G. R. Luevano, T. Wang, S. Özbayat, T. Michalka, and J. L. Drewniak, "Fast algorithm for minimizing the number of decap in power distribution networks," *IEEE Transactions on Electromagnetic Compatibility*, vol. 60, no. 3, pp. 725–732, June 2018, doi: 10.1109/TEMC.2017.2746677.
- [7] H. Su, S. S. Sapatnekar, and S. R. Nassif, "Optimal decoupling capacitor sizing and placement for standard-cell layout designs," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 22, no. 4, pp. 428–436, 2003, doi: 10.1109/TCAD.2003.809658.
- [8] S. M. Vazgen, H. S. Karo, V. A. Avetisyan, and A. T. Hakhverdyan, "On-chip decoupling capacitor optimization technique," 2017 IEEE 37th International Conference on Electronics and Nanotechnology (ELNANO), 2017, pp. 116–118, doi: 10.1109/ELNANO.2017.7939729.
- [9] J. Xu et al., "A novel system-level power integrity transient analysis methodology using simplified CPM model, physics-based equivalent circuit PDN model and small signal VRM model," in 2019 IEEE International Symposium on Electromagnetic Compatibility, Signal and Power Integrity (EMC+SIPI), 2019, pp. 205–210, doi: 10.1109/ISEMC.2019.8825256.
- [10] B. Zhao et al., "Systematic power integrity analysis based on inductance decomposition in a multi-layered PCB PDN," IEEE Electromagnetic Compatibility Magazine, vol. 9, no. 4, pp. 80–90, 2020, doi: 10.1109/MEMC.2020.9327998.
- [11] B. Zhao et al., "Physics-based circuit modeling methodology for system power integrity analysis and design," IEEE Transactions on Electromagnetic Compatibility, vol. 62, no. 4, pp. 1266–1277, Aug. 2020, doi: 10.1109/TEMC.2019.2927742.
- [12] I. Erdin and R. Achar, "Fast power integrity analysis of PDNs with arbitrarily shaped power-ground plane pairs," in 2020 IEEE Electrical Design of Advanced Packaging and Systems (EDAPS), 2020, pp. 1–3, doi: 10.1109/EDAPS50281.2020.9312920.
- [13] A. Cabani, P. Zhang, R. Khemmar, and J. Xu, "Enhancement of energy consumption estimation for electric vehicles by using machine learning," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 10, no. 1, pp. 215–223, 2021, doi: 10.11591/ijai.v10i1.pp215-223.
- [14] H. Ohmaid, S. Eddarouich, A. Bourouhou, and M. Timouyas, "Comparison between SVM and KNN classifiers for iris recognition using a new unsupervised neural approach in segmentation," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 3, pp. 429–438, 2020, doi: 10.11591/ijai.v9.i3.pp429-438.
- [15] T. Nguyen et al., "Comparative study of surrogate modeling methods for signal integrity and microwave circuit applications," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 11, no. 9, pp.

1369–1379, 2021, doi: 10.1109/TCPMT.2021.3098666.

- [16] C. H. Goay, N. S. Ahmad, and P. Goh, "Transient simulations of high-speed channels using CNN-LSTM with an adaptive successive halving algorithm for automated hyperparameter optimizations," *IEEE Access*, vol. 9, pp. 127644–127663, 2021, doi: 10.1109/ACCESS.2021.3112134.
- [17] K. S. Ooi, C. L. Kong, C. H. Goay, N. S. Ahmad, and P. Goh, "Crosstalk modeling in high-speed transmission lines by multilayer perceptron neural networks," *Neural Computing and Applications*, vol. 32, no. 11, pp. 7311–7320, 2020, doi: 10.1007/s00521-019-04252-3.
- [18] C. K. Ku, C. H. Goay, N. S. Ahmad, and P. Goh, "Jitter decomposition of high-speed data signals from jitter histograms with a pole-residue representation using multilayer perceptron neural networks," *IEEE Transactions on Electromagnetic Compatibility*, vol. 62, no. 5, pp. 2227–2237, Oct. 2020, doi: 10.1109/TEMC.2019.2936000.
- [19] M. Swaminathan, H. M. Torun, H. Yu, J. A. Hejase, and W. D. Becker, "Demystifying machine learning for signal and power integrity problems in packaging," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 10, no. 8, pp. 1276–1295, Aug. 2020, doi: 10.1109/TCPMT.2020.3011910.
- [20] F. D. J. Leal-Romo, J. L. Chávez-Hurtado, and J. E. Rayas-Sánchez, "Selecting surrogate-based modeling techniques for power integrity analysis," in 2018 IEEE MTT-S Latin America Microwave Conference (LAMC 2018), 2018, pp. 1–3, doi: 10.1109/LAMC.2018.8699021.
- [21] K. T. Chang et al., "Ultra high density IO fan-out design optimization with signal integrity and power integrity," in 2019 IEEE 69th Electronic Components and Technology Conference (ECTC), 2019, pp. 41–46, doi: 10.1109/ECTC.2019.00014.
- [22] S. Piersanti, R. Cecchetti, C. Olivieri, F. de Paulis, A. Orlandi, and M. Buecker, "Decoupling capacitors placement at board level adopting a nature-inspired algorithm," *Electronics*, vol. 8, no. 7, 2019, Art. no. 737, doi: 10.3390/electronics8070737.
- [23] K. Bharath, E. Engin, and M. Swaminathan, "Automatic package and board decoupling capacitor placement using genetic algorithms and M-FDM," in 2008 45th ACM/IEEE Design Automation Conference, 2008, pp. 560–565, doi: 10.1145/1391469.1391611.
- [24] H. Park et al., "Reinforcement learning-based optimal on-board decoupling capacitor design method," in 2018 IEEE 27th Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS), 2018, pp. 213–215, doi: 10.1109/EPEPS.2018.8534195.
- [25] H. Park et al., "Deep reinforcement learning-based optimal decoupling capacitor design method for silicon interposerbased 2.5-D/3-D ICs," in *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 10, no. 3, pp. 467–478, 2020, doi: 10.1109/TCPMT.2020.2972019.
- [26] L. Zhang et al., "Decoupling capacitor selection algorithm for PDN based on deep reinforcement learning," 2019 IEEE International Symposium on Electromagnetic Compatibility, Signal and Power Integrity (EMC+SIPI), 2019, pp. 616–620, doi: 10.1109/ISEMC.2019.8825249.
- [27] O. W. Bhatti and M. Swaminathan, "Impedance response extrapolation of power delivery networks using recurrent neural networks," in 2019 IEEE 28th Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS), 2019, pp. 1–3, doi: 10.1109/EPEPS47316.2019.193198.

BIOGRAPHIES OF AUTHORS



Yeong Kang Liew (D) (S) (B) received the B.Eng. in Electrical and Electronic Engineering from the Asian Institute of Medicine, Science and Technology (AIMST) University in 2018 and the M.Sc. degree in Electronic Systems Design Engineering from Universiti Sains Malaysia (USM) in 2020. His research interests are in the fields of electronics, signal and power integrity, and machine learning. He is currently a hardware design engineer in Intel Microelectronic (M) Sdn Bhd. He can be contacted at email: yeong.kang.liew@intel.com.



Nur Syazreen Ahmad **1** Preceived the B.Eng. (Hons) degree in Electrical and Electronic Engineering from the University of Manchester, United Kingdom in 2009, and Ph.D. degree in Control Systems from the same university in 2012. She is currently with the School of Electrical and Electronic Engineering, University Sains Malaysia. Her current research interest revolves around robust constrained control, intelligent control systems, computer-based control and autonomous mobile systems in wireless sensor network. She can be contacted at email: syazreen@usm.my.

Machine learning modeling of power delivery networks with varying decoupling ... (Yeong Kang Liew)





Patrick Goh B S S S P received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana-Champaign, Urbana, IL, USA in 2007, 2009, and 2012 respectively. Since 2012, he has been with the School of Electrical and Electronic Engineering, Universiti Sains Malaysia, where he currently specializes in the study of signal integrity for high-speed digital designs. His research interest includes the development of circuit simulation algorithms for computer-aided design tools. He was a recipient of the Raj Mittra Award in 2012 and the Harold L. Olesen Award in 2010, and has served on the technical program committee and international program committee in various IEEE and non-IEEE conferences around the world. He can be contacted at email: eep-atrick@usm.my.