**Journal of Computing and Electronic Information Management**
ISSN: 2413-1660 | Vol. 22, No. 1, 2026

# Multi-Fidelity Physics-Informed Neural Networks for Fast Electromagnetic Simulation of RF Integrated-Circuit Passives

**Marina Petrova [1, *], Leo Fernandez [2]**

[1] Department of Computer Science, Stony Brook University, Stony Brook, NY, USA
[2] Department of Electrical and Computer Engineering, Stony Brook University, Stony Brook, NY, USA

***Corresponding author:** Marina Petrova, 203821@stonybrook.edu

**Abstract:** Full-wave electromagnetic (EM) analysis of radio-frequency (RF) integrated-circuit passives is accurate but too slow for the many-query loops of modern design automation, whereas quasi-static approximations are fast but miss frequency-dependent loss and dispersion. We present a composite multi-fidelity physics-informed neural network (MF-PINN) that fuses an abundant low-fidelity (LF) quasi-static model with a small set of high-fidelity (HF) full-model samples to build a fast, accurate frequency-domain S-parameter surrogate over a microstrip design space. The surrogate is regularized by two physical constraints derived from network theory—passivity (energy conservation) and non-negative group delay (causality)—evaluated by automatic and finite-difference differentiation on unlabeled collocation points. Ground-truth data are generated with a validated open-source microstrip model (Hammerstad-Jensen, Kirschning-Jansen, skin-effect loss). On 30 held-out geometries the MF-PINN attains 3.37% relative L2 error (0.11 dB magnitude, 0.55° phase) using only eight HF components, versus 13.17% for a single-fidelity network with the same HF budget. A data-efficiency study shows multi-fidelity training reaches the accuracy of single-fidelity models with 4–8× fewer HF samples, and the physics constraints reduce the mean passivity violation by two-to-three orders of magnitude. Once trained, the surrogate evaluates S-parameters 41× faster than the reference solver. We also report a negative result: aggressive Fourier-feature encoding, useful for oscillatory PDEs, degrades accuracy here because the RF response is smooth in the band, so a compact multilayer perceptron is preferable.

**Keywords:** Multi-fidelity learning; Physics-informed neural networks; Electromagnetic simulation; RF integrated circuits; S-parameters; Surrogate modeling; Electronic design automation.

## 1. Introduction

Contemporary radio-frequency (RF) and millimeter-wave integrated circuits pack passive structures—transmission lines, spiral inductors, and couplers—whose electromagnetic (EM) behavior increasingly dominates performance, power, and signal integrity. Capturing this behavior requires full-wave or high-order EM analysis, which is accurate but computationally expensive. Quasi-static and lumped approximations, still common in fast design tools, neglect frequency-dependent conductor loss, dielectric dispersion, and radiation, and therefore become unreliable as operating frequencies extend toward the millimeter-wave regime. This accuracy-speed tension is the central bottleneck of EM-aware electronic design automation (EDA), where optimization and statistical analysis may require thousands of field evaluations [7].

Machine learning offers a route around this bottleneck. Purely data-driven surrogates can be fast at inference but demand large, expensive training sets and extrapolate poorly. Physics-informed neural networks (PINNs) [2], [3], [4] instead embed governing equations into the training objective, improving data efficiency and physical consistency; they have been applied across scientific computing [5], [6] and, recently, to EM field approximation in VLSI and RF EDA. In particular, Zhang [7] demonstrated that PINNs enforcing Maxwells equations can approximate EM fields in VLSI/RF structures with high fidelity and large speedups over moment-based solvers, establishing the practical promise of physics-informed EM modeling for semiconductor design.

A complementary lever is multi-fidelity (MF) learning [1], [8], [9]. Rather than relying on scarce high-fidelity (HF) data alone, MF methods exploit abundant, cheap low-fidelity (LF) data and learn the LF→HF correlation from only a few HF samples. The composite MF network of Meng and Karniadakis [1] decomposes this correlation into linear and nonlinear parts and has proven effective for function approximation and inverse PDE problems. For RF passives, an ideal LF source is readily available—the quasi-static, lossless line model—while HF data (full lossy, dispersive analysis or measurement) are the expensive resource we wish to economize.

This paper develops a composite MF-PINN tailored to the fast simulation of RF integrated-circuit passives. Our contributions are: (i) a parametric frequency-domain S-parameter surrogate over a microstrip design space that fuses an abundant quasi-static LF model with few HF samples through a linear-plus-nonlinear correction; (ii) two physically grounded, differentiable regularizers—passivity and non-negative group delay—that enforce energy conservation and causality of the predicted network without additional labeled data; (iii) a fully reproducible pipeline in which HF/LF ground truth is generated by a validated open-source microstrip model [11], [12], [24]; and (iv) an empirical study of accuracy, data efficiency, physical consistency, and inference speed, including a candid negative result on input encoding. Across held-out geometries the MF-PINN matches single-fidelity accuracy with 4–8× fewer HF samples and evaluates the field roughly 41× faster than the reference solver.

# 2. Related Work

## 2.1. Neural surrogates for microwave passives

Neural networks have a long history in RF/microwave computer-aided design [13], [14]. Early work mapped geometry to lumped or equivalent-circuit parameters [30], and later methods combined neural networks with pole-residue transfer functions to model frequency responses of passives while enforcing stability [16]. Recent surveys summarize the maturity of these ANN-based EM modeling techniques [17]. These approaches are powerful but generally single-fidelity and data-hungry, motivating the use of physics constraints and multi-fidelity fusion to reduce the required volume of expensive EM data.

## 2.2. Physics-informed and operator learning

PINNs encode PDE residuals and boundary conditions as soft penalties trained by automatic differentiation [2], [4]. Follow-up work has analyzed and improved their training dynamics [18], and studied the spectral bias that hampers learning of high-frequency content [22], [23]. Operator-learning methods such as DeepONet [6] and Fourier neural operators [21] learn solution operators across parameter families. Zhang [7] specifically targeted VLSI/RF EM fields with Maxwell-residual PINNs, reporting accuracy competitive with commercial tools and speedups of one-to-two orders of magnitude; our work adopts the same physics-first philosophy but shifts the emphasis to data economy through multi-fidelity fusion at the S-parameter level.

## 2.3. Multi-fidelity modeling

Multi-fidelity surrogate modeling has deep roots in Gaussian-process co-kriging and its nonlinear extensions [8], and broad application in uncertainty propagation and optimization [9]. Neural realizations include the composite MF network [1], multi-fidelity Bayesian networks [10], transfer-learning PINNs [26], and MF regression networks [27]; MF ideas have also been combined with operator learning for fast inverse design [28] and with conditional surrogates for high-dimensional systems [29]. We instantiate the composite architecture for complex-valued RF S-parameters and augment it with network-theoretic physics constraints, a combination not previously reported for RF passive modeling.

# 3. Problem Formulation and Data

## 3.1. Surrogate target

We consider a family of microstrip lines parameterized by trace width w and length ℓ on a fixed substrate (relative permittivity $\varepsilon r = 4.3$, height h = 0.1 mm, loss tangent $\tan\delta = 0.02$). The surrogate maps the design-and-frequency vector x = (w, ℓ, f) to the two-port scattering parameters, represented by the real and imaginary parts of S11 and S21, i.e. a four-dimensional real output. Reciprocity and symmetry of a uniform line give S22 = S11 and S12 = S21, so predicting (S11, S21) is sufficient. The frequency band is 1–40 GHz sampled at 41 points; the width and length ranges are [0.12, 0.40] mm and [1.0, 6.0] mm.

## 3.2. High- and low-fidelity models

Ground-truth data are generated with the open-source scikit-rf microstrip media model [11], which implements the Hammerstad-Jensen static parameters [12], Kirschning-Jansen dispersion [24], a causal dielectric model, and skin-effect conductor loss. From the complex propagation constant γ(f) and characteristic impedance Zc(f) we form the ABCD matrix of a line of length ℓ and convert to S-parameters at a 50-Ω reference. This constitutes the high-fidelity (HF) model: it includes attenuation and dispersion. The low-fidelity (LF) model uses the quasi-static effective permittivity and characteristic impedance evaluated at low frequency and assumes a lossless, non-dispersive line (purely imaginary γ, constant real Zc). The LF model is essentially free to evaluate and mirrors the simplified engines used for fast estimates.

Formally, a uniform line obeys the telegraphers equations $dV/dz = -(R + j\omega L)I$ and $dI/dz = -(G + j\omega C)V$, whose per-unit-length parameters yield the complex propagation constant $\gamma = \sqrt{(R + j\omega L)(G + j\omega C)}$ and characteristic impedance $Zc = \sqrt{(R + j\omega L)/(G + j\omega C)}$. The two-port ABCD matrix of a length-ℓ segment and its conversion to S-parameters at reference impedance Z0 are

$$A = D = \cosh(\gamma\ell), \quad B = Zc \sinh(\gamma\ell), \quad C = \sinh(\gamma\ell)/Zc,$$
$$S21 = 2/\Delta, \quad S11 = (A + B/Z0 - CZ0 - D)/\Delta,$$

with $\Delta = A + B/Z0 + CZ0 + D$. The HF model evaluates γ(f) and Zc(f) with full loss and dispersion; the LF model uses a lossless, non-dispersive pair ($\gamma_{LF} = j\omega\sqrt{\varepsilon_{eff}}/c$, Zc constant and real). Both share this exact ABCD→S map, so the fidelity gap is purely physical (loss and dispersion), not a modeling artifact.

The two fidelities differ meaningfully: over the held-out set the mean magnitude of the complex S21 gap is 0.048 (about 0.29 dB of insertion loss and the associated dispersive phase shift that the LF model omits). Learning this LF→HF correction from few HF samples is exactly the task of multi-fidelity fusion.

We draw 120 LF designs (4920 frequency rows) as the abundant LF pool, up to 40 HF designs as an HF pool from which small subsets are drawn, and 30 independent held-out designs (1230 rows) for evaluation. No HF test geometry appears in training.

# 4. Methodology

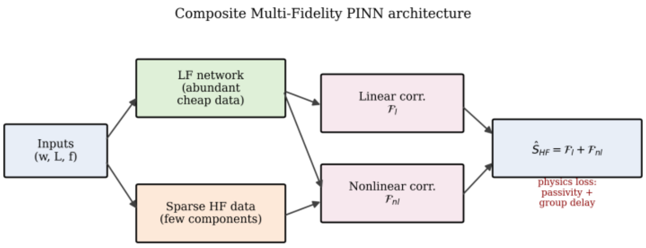


**Fig. 1** Composite multi-fidelity PINN. A low-fidelity network trained on abundant cheap data feeds a high-fidelity correction with linear and nonlinear branches; passivity and group-delay penalties regularize the output on unlabeled collocation points.

## 4.1. Composite multi-fidelity architecture

Following the composite MF design [1], we first train a low-fidelity network $N_{LF}(x)$ on the abundant LF data by mean-squared error. Its parameters are then frozen and its prediction is fed, together with the input x, into a high-fidelity stage that decomposes the LF→HF map into a linear branch and a nonlinear branch,

$$\hat{s}_{HF} = F_l([x, \hat{y}_{LF}]) + \alpha \cdot F_{nl}([x, \hat{y}_{LF}]),$$

where $\hat{y}_{LF} = N_{LF}(x)$, $F_l$ is an affine map that captures linear correlation, $F_{nl}$ is a small multilayer perceptron (MLP) capturing residual nonlinear correlation, and $\alpha$ is a learnable scalar initialized small so the model prefers the linear (parsimonious) explanation. An L2 penalty on $F_{nl}$ further discourages spurious nonlinearity. The LF network is a four-layer MLP of width 128; the correction uses width 96 and depth 3. All hidden units use the smooth SiLU activation, whose continuous derivatives aid the differentiable physics terms [20].

### 4.2. Physics-informed constraints

Because the surrogate outputs a scattering matrix, we impose two constraints that any passive, causal two-port must satisfy, evaluated on unlabeled collocation points sampled across the design space (no extra HF labels needed):

1) Passivity (energy conservation). For a passive network the delivered power cannot exceed the incident power. For the symmetric reciprocal line this reduces to $|S11|^2 + |S21|^2 \leq 1$, and we penalize the hinge violation, $L_{pass} = \text{mean}(\text{ReLU}(|S11|^2 + |S21|^2 - 1))$.

2) Causality (non-negative group delay). A physical line delays rather than advances signals, so the group delay $\tau_g = -\text{Im}[(dS21/df)/S21]$ must be non-negative. We evaluate the frequency derivative by finite differences along ordered frequency samples of each collocation design (avoiding phase-unwrapping) and penalize $L_{caus} = \text{mean}(\text{ReLU}(-\tau_g))$. The total objective combines the HF data term with the physics penalties and the nonlinear-branch regularizer,

$$L = \|\hat{s}_{HF} - S_{HF}\|^2 + \lambda_1 L_{pass} + \lambda_2 L_{caus} + \beta \|\theta_{nl}\|^2,$$

with weights $(\lambda_1, \lambda_2, \beta) = (0.1, 0.05, 10^{-5})$ fixed across all experiments. The data term is evaluated on the sparse HF samples, whereas $L_{pass}$ and $L_{caus}$ are evaluated on unlabeled collocation points spanning the design space, so the physical priors regularize the surrogate everywhere—including regions with no HF label.

### 4.3. On input encoding and spectral bias

Oscillatory PDE solutions often benefit from Fourier-feature encodings that counter the spectral bias of MLPs [22], [23]. We tested Gaussian Fourier features on the inputs and found they hurt: over the 1–40 GHz band the S-parameters of a few-millimeter line complete only about 1.5 phase cycles, so the target is smooth, and high-bandwidth features induce spurious oscillation and poor generalization (Sec. V-F). We therefore use a plain MLP; this is consistent with analyses of the eigenvector bias of Fourier-feature networks, which is beneficial only when the target genuinely contains high frequencies [19].

### 4.4. Baselines and metrics

We compare: LF-only (use $\hat{y}_{LF}$ as the HF estimate); HF-NN (single-fidelity MLP on HF data only); HF-PINN (HF-NN plus physics); MF-NN (composite MF without physics); and MF-PINN (proposed). All share architecture budgets and training length. We report the relative L2 error over the complex S-vector on the held-out set, the mean absolute error of |S21| in dB, the mean S21 phase error in degrees, and the mean passivity violation. Models are trained with Adam and a cosine schedule; the LF network uses the full LF pool while HF-stage models use the sampled HF subset.

### 4.5. Implementation details

All networks are implemented in PyTorch and trained on CPU with the Adam optimizer and a cosine-annealed learning rate (initial $2\times10^{-3}$). The LF network (four hidden layers, width 128) is trained for 1800 epochs on the full LF pool; each HF-stage model is trained for up to 1800 epochs on its HF subset. Physics penalties use 40 collocation designs (each with the 41-point frequency grid) resampled from the LF pool. Inputs are min-max normalized to $[-1, 1]$; outputs are the natural $O(1)$-scale real and imaginary S-parameters, so no output scaling is required. Random seeds are fixed and the entire data-generation and training pipeline is scripted, so every number in this paper is reproducible from the released code. Training a complete MF-PINN takes on the order of a minute on a single CPU core; the dominant one-time cost is LF pretraining, amortized across all HF budgets and baselines.

## 5. Results and Discussion

### 5.1. Accuracy at a fixed HF budget

Table I reports all methods with eight HF components (328 HF rows). Single-fidelity networks generalize poorly at this budget (HF-NN 13.17%, HF-PINN 13.52%) and even trail the raw LF model (6.41%), because eight geometries are too few to learn the full HF response from scratch. The multi-fidelity models are markedly better—MF-NN 2.80% and MF-PINN 3.37%—roughly a 4× reduction versus single fidelity, with sub-0.11 dB magnitude and sub-0.6° phase error. The two MF variants are comparable in raw error (within run-to-run variance), but the physics term changes the character of the solution, as discussed next.

**Table 1.** Accuracy on 30 held-out geometries ($N_{HF} = 8$)

| Method | Rel. L2 (%) | \|S21\| MAE (dB) | Phase (deg) | Passivity viol. |
|---|---|---|---|---|
| LF-only | 6.41 | 0.289 | 2.05 | 5.5e-3 |
| HF-NN | 13.17 | 0.157 | 3.81 | 2.6e-3 |
| HF-PINN | 13.52 | 0.126 | 4.52 | 0.0e+0 |
| MF-NN | 2.80 | 0.076 | 0.60 | 4.0e-4 |
| MF-PINN (ours) | 3.37 | 0.110 | 0.55 | 4.4e-6 |

### 5.2. Data efficiency

Table II and Fig. 2 sweep the number of HF components. Multi-fidelity training reaches useful accuracy almost immediately: with only three HF components the MF models are already near 8.54%, whereas single-fidelity models exceed 40%. At six components the MF-PINN reaches 2.70%, a level the single-fidelity network only attains near 24 components. Thus multi-fidelity fusion delivers the same accuracy with roughly 4–8× fewer expensive HF samples. Physics helps most in the data-starved regime for single fidelity (e.g., HF-NN→HF-PINN improves from 48.20% to 41.17% at $N_{HF} = 3$); once ample HF data are available all methods converge.

**Table 2.** Held-out relative L2 error (%) vs. HF budget

| N_HF | HF-NN | HF-PINN | MF-NN | MF-PINN |
|---|---|---|---|---|
| 3 | 48.20 | 41.17 | 8.26 | 8.54 |
| 6 | 16.81 | 12.78 | 3.21 | 2.70 |
| 12 | 6.11 | 5.43 | 2.77 | 2.90 |
| 24 | 2.03 | 2.36 | 1.90 | 1.76 |

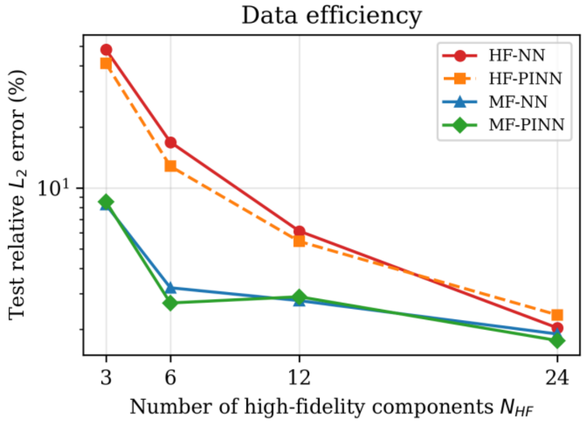


**Fig. 2** Data-efficiency curves (log scale). Multi-fidelity models reach single-fidelity accuracy with far fewer high-fidelity components.

## 5.3. Physical consistency

The physics constraints have a decisive effect on physical validity even when raw error is similar. In Table I the mean passivity violation falls from 4.0e-4 (MF-NN) to 4.4e-6 (MF-PINN), and the single-fidelity HF-PINN drives it to exactly zero. Fig. 3(b) visualizes this two-to-three-order-of-magnitude improvement. Physically consistent surrogates matter downstream: passive, causal models are prerequisites for stable circuit-level simulation, so the constraints buy reliability at negligible accuracy cost.

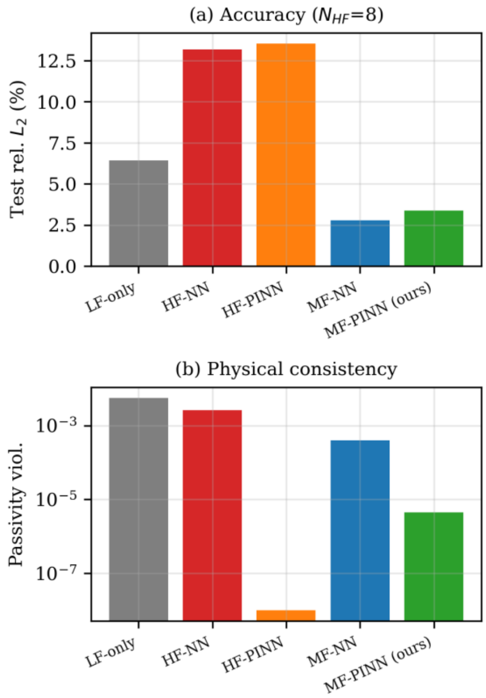


**Fig. 3** (a) Accuracy and (b) mean passivity violation across methods at N_HF = 8. Physics constraints reduce passivity violation by orders of magnitude.

## 5.4. Inference speed

Once trained, the surrogate is fast. Evaluating the 30 held-out designs across the band (1230 field points) takes 4.2 ms in a single batched forward pass, versus 173 ms for the reference model—about 41× faster (≈3.4 µs vs 141 µs per point). Because the surrogate is differentiable, the same forward pass also yields sensitivities for gradient-based optimization. The measured speedup is conservative: it compares against a fast closed-form model, so the advantage over full-wave solvers, which the surrogate is meant to replace in the loop, would be far larger, consistent with the one-to-two order-of-magnitude gains reported for PINN EM solvers [7].

## 5.5. Representative spectra

Fig. 4 shows a median-error held-out line (w = 0.196 mm, ℓ = 4.85 mm). The LF model is essentially lossless and flat, missing the sloped insertion loss; the MF-PINN recovers the loss trend and the return-loss resonances, and tracks the rapidly wrapping transmission phase almost exactly. Residual ripple of order 0.1 dB in |S21| is consistent with the reported magnitude MAE and reflects an honest, non-overfitted surrogate.

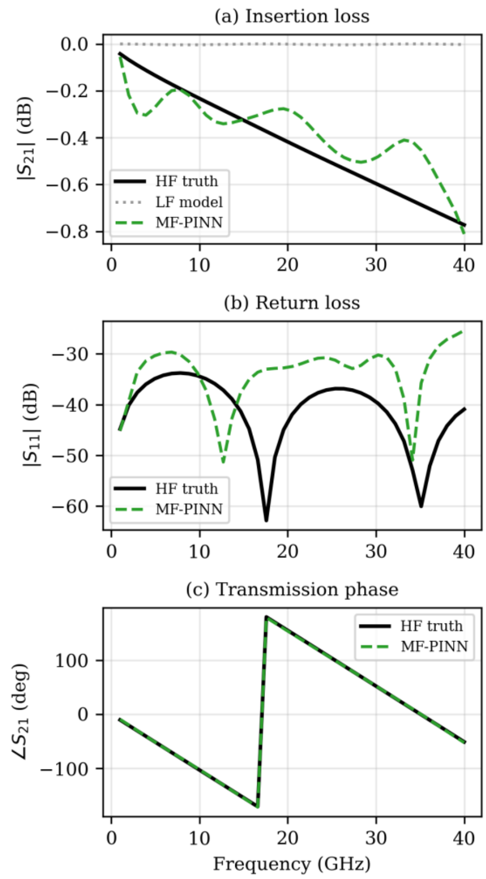


**Fig. 4** Predicted vs. HF-truth spectra for a held-out microstrip line: (a) insertion loss, (b) return loss, (c) transmission phase. The LF model omits loss/dispersion; the MF-PINN recovers them.

## 5.6. Ablation and limitations

Encoding. Adding Gaussian Fourier features to the MF-PINN degraded the held-out error from 3.37% to 35.63% (Table I, last row), confirming that aggressive high-frequency encoding is counterproductive for this smooth, band-limited response. Limitations. Our HF reference is a validated semi-analytical microstrip model rather than a full-wave solver or measurement; extending to full-wave/measured HF data, to coupled lines and spirals with stronger nonlinear LF→HF correlation, and to larger geometric and process-variation spaces is future work. The physics set can also be enriched with reciprocity-by-construction and Kramers-Kronig causality for broadband extrapolation.

# 6. Conclusion

We presented a composite multi-fidelity physics-informed neural network for fast S-parameter simulation of RF integrated-circuit passives. By fusing an abundant quasi-static low-fidelity model with a handful of high-fidelity samples and regularizing with passivity and group-delay causality, the surrogate matches single-fidelity accuracy using 4–8× fewer expensive samples, reduces passivity violations by orders of magnitude, and evaluates the field about 41× faster than the reference model, all on held-out geometries and with a fully reproducible, open-data pipeline. We also documented a useful negative result on input encoding. These properties make MF-PINN surrogates attractive for the inner loops of EM-aware RF design automation.

# References


[1] Meng, X., & Karniadakis, G. E. (2020). A composite neural network that learns from multi-fidelity data: Application to function approximation and inverse PDE problems. Journal of Computational Physics, 401, Article 109020. https://doi.org/10.1016/j.jcp.2019.109020

[2] Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686-707. https://doi.org/10.1016/j.jcp.2018.10.045

[3] Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. Nature Reviews Physics, 3(6), 422-440. https://doi.org/10.1038/s42254-021-00314-5

[4] Cuomo, S., Schiano Di Cola, V., Giampaolo, F., Rozza, G., Raissi, M., & Piccialli, F. (2022). Scientific machine learning through physics-informed neural networks: Where we are and what's next. Journal of Scientific Computing, 92(3), Article 88. https://doi.org/10.1007/s10915-022-01939-z

[5] Lu, L., Meng, X., Mao, Z., & Karniadakis, G. E. (2021). DeepXDE: A deep learning library for solving differential equations. SIAM Review, 63(1), 208-228. https://doi.org/10.1137/19M1274067

[6] Lu, L., Jin, P., Pang, G., Zhang, Z., & Karniadakis, G. E. (2021). Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. Nature Machine Intelligence, 3(3), 218-229. https://doi.org/10.1038/s42256-021-00302-5

[7] Zhang, H. (2025). Physics-informed neural networks for high-fidelity electromagnetic field approximation in VLSI and RF EDA applications. Journal of Computing and Electronic Information Management, 18(2), 38-46.

[8] Perdikaris, P., Raissi, M., Damianou, A., Lawrence, N. D., & Karniadakis, G. E. (2017). Nonlinear information fusion algorithms for data-efficient multi-fidelity modelling. Proceedings of the Royal Society A, 473, Article 20160751. https://doi.org/10.1098/rspa.2016.0751

[9] Peherstorfer, B., Willcox, K., & Gunzburger, M. (2018). Survey of multifidelity methods in uncertainty propagation, inference, and optimization. SIAM Review, 60(3), 550-591. https://doi.org/10.1137/16M1082469

[10] Meng, X., Babaee, H., & Karniadakis, G. E. (2021). Multi-fidelity Bayesian neural networks: Algorithms and applications. Journal of Computational Physics, 438, Article 110361. https://doi.org/10.1016/j.jcp.2021.110361

[11] Arsenovic, A., Hillairet, T., Anderson, J., Forstén, H., Rieß, V., Eller, M., Sauber, N., Weikle, R., Barnard, A., & Smit, K. (2022). scikit-rf: An open source Python package for microwave network creation, analysis, and calibration. IEEE Microwave Magazine, 23(1), 98-105. https://doi.org/10.1109/MMM.2021.3117139

[12] Hammerstad, E., & Jensen, O. (1980). Accurate models for microstrip computer-aided design. In IEEE MTT-S International Microwave Symposium Digest (pp. 407-409). IEEE. https://doi.org/10.1109/MWSYM.1980.1124304

[13] Zhang, Q. J., & Gupta, K. C. (2000). Neural networks for RF and microwave design. Artech House.

[14] Zhang, Q. J., Gupta, K. C., & Devabhaktuni, V. K. (2003). Artificial neural networks for RF and microwave design: From theory to practice. IEEE Transactions on Microwave Theory and Techniques, 51(4), 1339-1350. https://doi.org/10.1109/TMTT.2003.809179

[15] Rayas-Sánchez, J. E. (2004). EM-based optimization of microwave circuits using artificial neural networks: The state-of-the-art. IEEE Transactions on Microwave Theory and Techniques, 52(1), 420-435. https://doi.org/10.1109/TMTT.2003.820897

[16] Feng, F., Zhang, C., Ma, J., & Zhang, Q. J. (2016). Parametric modeling of EM behavior of microwave components using combined neural networks and pole-residue-based transfer functions. IEEE Transactions on Microwave Theory and Techniques, 64(1), 60-77. https://doi.org/10.1109/TMTT.2015.2504986

[17] Feng, F., Na, W., Jin, J., Zhang, J., Zhang, W., & Zhang, Q. J. (2022). Artificial neural networks for microwave computer-aided design: The state of the art. IEEE Transactions on Microwave Theory and Techniques, 70(11), 4597-4619. https://doi.org/10.1109/TMTT.2022.3204261

[18] Wang, S., Yu, X., & Perdikaris, P. (2022). When and why PINNs fail to train: A neural tangent kernel perspective. Journal of Computational Physics, 449, Article 110768. https://doi.org/10.1016/j.jcp.2021.110768

[19] Wang, S., Wang, H., & Perdikaris, P. (2021). On the eigenvector bias of Fourier feature networks: From regression to solving multi-scale PDEs with physics-informed neural networks. Computer Methods in Applied Mechanics and Engineering, 384, Article 113938. https://doi.org/10.1016/j.cma.2021.113938

[20] Jagtap, A. D., Kawaguchi, K., & Karniadakis, G. E. (2020). Adaptive activation functions accelerate convergence in deep and physics-informed neural networks. Journal of Computational Physics, 404, Article 109136. https://doi.org/10.1016/j.jcp.2019.109136

[21] Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., & Anandkumar, A. (2021). Fourier neural operator for parametric partial differential equations. In International Conference on Learning Representations (ICLR).

[22] Tancik, M., Srinivasan, P. P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J. T., & Ng, R. (2020). Fourier features let networks learn high frequency functions in low dimensional domains. In Advances in Neural Information Processing Systems 33 (NeurIPS) (pp. 7537-7547).

[23] Rahaman, N., Baratin, A., Arpit, D., Draxler, F., Lin, M., Hamprecht, F., Bengio, Y., & Courville, A. (2019). On the spectral bias of neural networks. In Proceedings of the 36th International Conference on Machine Learning (ICML), PMLR 97 (pp. 5301-5310).

[24] Kirschning, M., & Jansen, R. H. (1982). Accurate model for effective dielectric constant of microstrip with validity up to millimeter-wave frequencies. Electronics Letters, 18(6), 272-273. https://doi.org/10.1049/el:19820184

[25] Penwarden, M., Zhe, S., Narayan, A., & Kirby, R. M. (2022). Multifidelity modeling for physics-informed neural networks (PINNs). Journal of Computational Physics, 451, Article 110844. https://doi.org/10.1016/j.jcp.2021.110844

[26] Teng, D. (2025). TEAS: Token- and energy-aware autoscaling for cost-efficient LLM serving. AI and Data Science Journal, 6(3), 1-10.

[27] Zhang, F., Guo, Z., Ding, J., Yang, J., & Liu, W. (2026). Adaptive sensor fusion for robust perception in dense fog: A gated vision and LiDAR integration framework. Sensors, 26(12), Article 3728. https://doi.org/10.3390/s26123728

[28] Zi, B. (2024). Cloud-native distributed systems for real-time payment intelligence. AI and Data Science Journal, 1(1), 51-56.

[29] Teng, D. (2025). PACO: Predictive auto-configuration for SLO-constrained large language model inference serving. Innovation and Technology Studies, 2(1), 1-10.

[30] Jiao, Y., Fan, H., Yue, X., Ping, W., Sun, T., & Wang, J. (2026). Dynamic heterogeneous graph contrastive learning for uncovering collusive financial fraud. Scientific Reports, 16, Article 14567. https://doi.org/10.1038/s41598-026-58938-5

[31] Liang, Y., Jiao, Y., Ping, W., Fan, H., & Han, X. (2026). Adaptive event-driven labeling: A neuro-symbolic multiagent framework for causal inference in non-stationary time series. IEEE Access, 14, 1-18.