DATE 25
DESIGN, AUTOMATION AND TEST IN EUROPE
THE EUROPEAN EVENT FOR ELECTRONIC SYSTEM DESIGN & TEST
31 MARCH – 2 APRIL 2025
LYON, FRANCE
CENTRE DE CONGRÈS DE LYON

IR-Fusion: A Fusion Framework for Static IR Drop Analysis Combining Numerical Solution and Machine Learning

Feng Guo[1], Jianwang Zhai[1], Jingyu Jia[1], Jiawei Liu[1], Kang Zhao[1], Bei Yu[2], Chuan Shi[1]

[1] Beijing University of Posts and Telecommunications
[2] The Chinese University of Hong Kong

Apr. 1, 2025

Outline

Introduction

Background: Power Grid and IR Drop Analysis

- The on-chip power grid (PG) transfers voltage and current to each working cell, and IR drop analysis involves obtaining the voltage drop caused by parasitics between the power pads and cells.

1. Ensuring the worst-case IR drop values are within specified limits is essential.
2. IR drop analysis becomes very **time-consuming** in industrial-scale designs using **traditional analysis methods**.
3. PG is designed from the top-level metal layer, which is connected to the power supplier, down through inter-layer vias, and finally to the active cells.

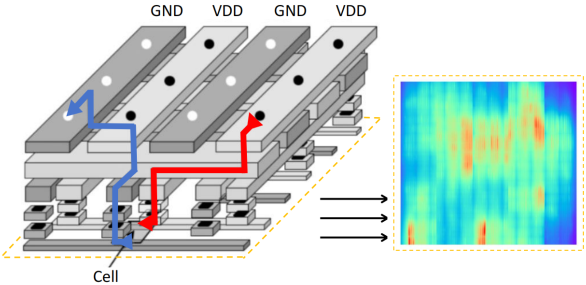


Fig. 1 The multiple-layer structure of PG.

- Many numerical methods have been proposed for this process, including direct solvers[1,2], iterative solvers[3], and other specialized solvers[4].
- The system matrix of a n-node PG network can be formulated as a linear system:

$$\boldsymbol{G}\boldsymbol{x} = \boldsymbol{I} \tag{1}$$

- As the number of nodes in the PG grows exponentially, traditional methods struggle with **longer solution times** or even become infeasible due to high computational demands and memory demands.
- Consequently, the necessity for ML methods becomes evident.

---

[1] T. A. Davis, et al. (2010). "Algorithm 907: KLU, a direct sparse solver for circuit simulation problems," in *Article TOMS*, pp. 1–17.

[2] Y. Chen, et al. (2008). "Algorithm 887: CHOLMOD, supernodal sparse Cholesky factorization and update/downdate," in *Article TOMS*, pp. 1–14.

[3] T.-H. Chen, et al. (2001). "Efficient large-scale power grid analysis based on preconditioned Krylov-subspace iterative methods," in *Proc. DAC*, pp. 559–562.

[4] Z. Liu, et al. (2024). "PowerRChol: Efficient Power Grid Analysis Based on Fast Randomized Cholesky Factorization," in *Proc. DAC*, pp. 1–6.

Background: Machine Learning for IR Drop Analysis

- To address inefficiencies, machine learning (ML)-based methods have been proposed as a promising alternative for accelerating IR drop analysis:

  1. IREDGe[5].
  2. MAVREC[6]
  3. PGAU[7]
  4. MAUnet[8]

- They still face the problem of **insufficiently fine modeling granularity**.
- They struggle with issues related to model **interpretability and generalizability**, which can limit their adoption in practical design environments.

[5] V. A. Chhabria, et al. (2021). "Thermal and IR drop analysis using convolutional encoder-decoder networks," in *Proc. ASP-DAC*, pp. 690–696.

[6] V. A. Chhabria, et al. (2021). "MAVIREC: ML-aided vectored IR-drop estimation and classification," in *Proc. DATE*, pp. 1825–1828.

[7] F. Guo, et al. (2024). "PGAU: Static IR Drop Analysis for Power Grid using Attention U-Net Architecture and Label Distribution Smoothin," in *Proc. GLSVLSI*, pp. 452–458

[8] M. Wang, et al. (2022). "MAUnet: Multiscale attention U-Net for effective IR drop prediction," in *Proc. DAC*, pp. 1–6.

Key Point: Integrating Numerical and ML Methods

- **Can numerical and ML methods be combined for a better trade-off in speed, accuracy, and scalability? Yes!**
- Most numerical methods solve large-scale linear systems iteratively, where more iterations yield greater accuracy but require longer runtime.
- By integrating ML, we can perform fewer iterations to **obtain a rough solution and refine it using ML**.
- This fusion enables a better understanding of complex physical or geometric systems, while offering more fine-grained and efficient modeling.

Preliminaries

Problem Formulation

- This work aims to **fuse the numerical method with ML** to achieve better performance in static IR drop analysis while **focusing on the IR drop of the cell at the bottom layer**.
- Each feature map, denoted as $\boldsymbol{P}_{\mathrm{map}i}$ essentially serves as a spatial representation of the inherent properties of the PG.
- The IR drop of working cells in the entire PG is also converted to a data matrix and represented as y. The algorithm F tries to give the closest prediction $F*$ based on all the input features $\left(\boldsymbol{P}_{\mathrm{map}_1}, \ldots, \boldsymbol{P}_{\mathrm{map}_n}\right)$, formulated as:

$$F^* = \arg\min \mathrm{Loss}\left(F\left(\boldsymbol{P}_{\mathrm{map}_1}, \ldots, \boldsymbol{P}_{\mathrm{map}_n}\right), y\right). \tag{2}$$

Methodologies

Overall Framework: IR-Fusion

- IR-Fusion consists of several components:
  1. An efficient AMG-PCG solver
  2. Hierarchical numerical-structural fusion
  3. Inception Attention U-Net model
  4. Augmented curriculum learning

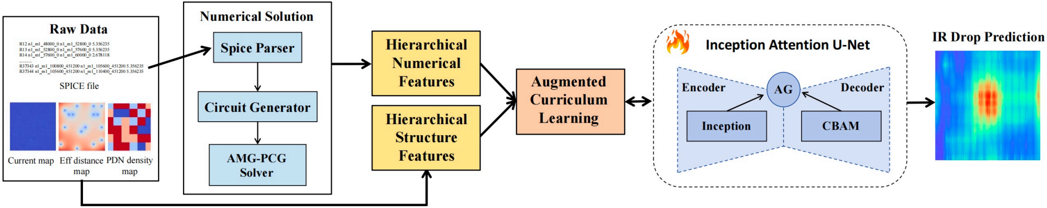


Fig. 2 The overview of IR-Fusion, a fusion framework for static IR drop analysis combining numerical solution and ML.

Step 1: Numerical Solution using AMG-PCG[9]

- In the numerical solution phase
  1. A spice parser
  2. A circuit generator
  3. The algebraic multigrid preconditioned conjugate gradient (AMG-PCG) method in PowerRush

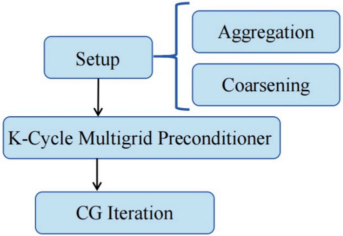


Fig. 3 The illustration of AMG-PCG solver.

[9] J. Yang, et al. (2013). "PowerRush: An efficient simulator for static power grid analysis", in *Article. TLVSI*, pp. 2103–2116.

Step 1: Numerical Solution using AMG-PCG

- Setup Stage Transfer operators connecting fine and coarse levels are represented by the preconditioning matrix $\boldsymbol{M}^{-1}$, which ensures the residual $\boldsymbol{r}_k = \boldsymbol{I} - \boldsymbol{G}\boldsymbol{x}_k$ and search direction $\boldsymbol{p}_k$ are accurately transferred between grids.
- Preconditioning Phase

$$\boldsymbol{r}_{k+1} = \boldsymbol{r}_k - \frac{\boldsymbol{r}_k^\top \boldsymbol{M}^{-1}\boldsymbol{r}_k}{\boldsymbol{p}_k^\top \boldsymbol{G}\boldsymbol{p}_k}\boldsymbol{G}\boldsymbol{p}_k,$$

where $\boldsymbol{M}^{-1}\boldsymbol{r}_k$ represents the correction on multiple grid levels, ensuring fast convergence by addressing errors at various scales. This correction accelerates the reduction of the residual in each iteration.

- CG Method

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \frac{\boldsymbol{r}_k^\top \boldsymbol{M}^{-1}\boldsymbol{r}_k}{\boldsymbol{p}_k^\top \boldsymbol{G}\boldsymbol{p}_k}\boldsymbol{p}_k, \tag{3}$$

$$\boldsymbol{p}_{k+1} = \boldsymbol{M}^{-1}\boldsymbol{r}_{k+1} + \frac{\boldsymbol{r}_{k+1}^\top \boldsymbol{M}^{-1}\boldsymbol{r}_{k+1}}{\boldsymbol{r}_k^\top \boldsymbol{M}^{-1}\boldsymbol{r}_k}\boldsymbol{p}_k. \tag{4}$$

This ensures that the search direction is adjusted based on the multilevel corrections, allowing the CG method to converge more quickly to the final solution $\boldsymbol{x}$.

Step 2: Hierarchical Numerical-Structural Information Fusion

- Based on the row w and height l from Library Exchange Format (LEF), a design's layer of size $W_c \times L_c$ translates to an image of $W\,(= W_c//w) \times L\,(= L_c//l)$ pixels.
- Each metal layer corresponds to a generated feature map, allowing the PG to produce feature maps that align with the same number of grid layers in total.

Features

Given the limited representation of designs, our method extracts more hierarchical structure features using the PG spice file and cell layer features:

1. **The current map** for each layer, representing the current distribution, is allocated proportionally based on the contribution from each layer, which is tied to resistance.
2. **The effective distance**, calculated as the reciprocal of the sum of the reciprocals of Euclidean distances, measures proximity to voltage sources.
3. **The PDN density map** is derived from the average PDN pitch within each grid as detailed in the spice file.
4. **The resistance** and **shortest path resistance maps** are also computed based on their physical significance.

- Based on PGAU, we design our Inception Attention Unet.
- **The inception module** is applied to enhance the network's ability to capture both local details and broader context.
- **The convolutional block attention module (CBAM)** is incorporated to focus on various scales and directions in subsequent decoder stages.

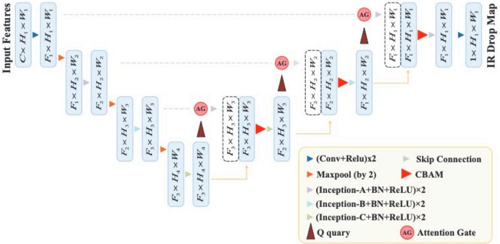


Fig. 4 The architecture of Inception Attention U-Net.

Step 4: Augmented Curriculum Learning

- We augment the training data by applying various transformations to each image-based input.
- We apply predefined CL for the train set.

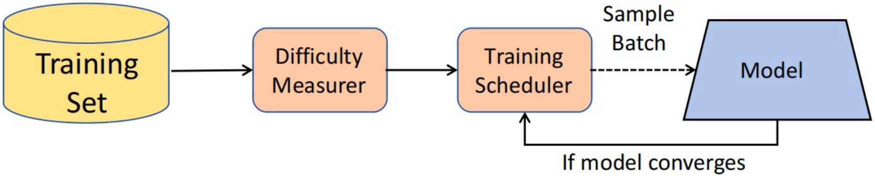


Fig. 5 The illustration of predefined curriculum learning.

Evaluation

Baselines

- Baselines:
  1. IREDGe[10].
  2. MAVREC[11]
  3. IRPnet[12]
  4. PGAU[13]
  5. MAUnet[14]
  6. Contest Winner[15]

---

[10] V. A. Chhabria, et al. (2021). "Thermal and IR drop analysis using convolutional encoder-decoder networks," in *Proc. ASP-DAC*, pp. 690–696.

[11] V. A. Chhabria, et al. (2021). "MAVIREC: ML-aided vectored IR-drop estimation and classification," in *Proc. DATE*, pp. 1825–1828.

[12] Y. Meng, et al. (2024). "Circuits physics constrained predictor of static IR drop with limited data," in *Proc. DATE*, pp. 1-2.

[13] F. Guo, et al. (2024). "PGAU: Static IR Drop Analysis for Power Grid using Attention U-Net Architecture and Label Distribution Smoothin," in *Proc. GLSVLSI*, pp. 452–458

[14] M. Wang, et al. (2022). "MAUnet: Multiscale attention U-Net for effective IR drop prediction," in *Proc. DAC*, pp. 1–6.

[15] Winners at ICCAD 2023 Contest. [Online]. Available: https://www.iccad-contest.org/2023/Winners.html.

Datasets

- The ICCAD2023 dataset[16], specialized for the static IR drop prediction task, is used for evaluation. It contains **120 designs, 20 of which are real designs, and the rest were artificially generated** based on BeGAN[17], named fake designs, close to realistic PGs. We perform the following setup on the dataset:

1. We follow the contest setup, using 10 real designs for testing and the rest for training.
2. Data augmentation increases the dataset size fourfold, with oversampling applied: fake designs are doubled, and real ones are quintupled.
3. Following a curriculum learning strategy, fake designs are categorized as "easier," while real designs are classified as "harder."

[16] Winners at ICCAD 2023 Contest. [Online]. Available: https://www.iccad-contest.org/2023/Winners.html.

[17] V. A. Chhabria, et al. (2021). "BeGAN: Power grid benchmark generation using a process-portable GAN-based methodology," in *Proc. ICCAD*, pp. 1–8.

Metrics

- Metrics:

  1. **Mean absolute error (MAE)** shows the average of the absolute difference between a prediction and the ground truth..
  2. **The F1 score** reflects the accuracy and comprehensiveness of the prediction for the hotspots region, fomulated as:

  $$P = \frac{TP}{TP + FP}, R = \frac{TP}{TP + FN}, F1 = \frac{2 \times P \times R}{P + R}. \tag{5}$$

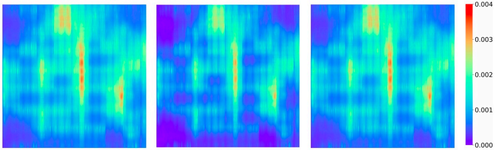
  3. **MIRDE** is the error in the region of maximum IR drop. Since designers are more concerned with the worst-case of IR drop, its modeling error is extremely critical.
  4. **Runtime**

Experiment 1: Main Experiment with Baselines

- IR-Fusion achieves better performance with the improvement of 28.3% on MAE, 14.5% on F1, and 27.6% on MIRDE, with no significant time cost increase. compared to the SOTA baseline, i.e., MAUnet.
- IR-Fusion still outperforms all baselines in MIRDE, representing more accuracy in the worst-case region.
- Our proposed fusion framework achieves more outstanding and robust performance within an acceptable runtime.

TABLE I Main results. The unit of MAE and MIRDE is $\times 10^{-4}V$ and runtime is s.

Methods	MAE↓	F1↑	Runtime↓	MIRDE↓
IREDGe [17]	3.75	0.49	1.55	7.52
MAVIREC [18]	2.78	0.46	1.97	5.88
IRPnet [19]	1.66	0.61	2.22	5.25
PGAU [20]	1.72	0.60	2.07	5.02
MAUnet [21]	1.06	0.62	2.31	4.38
Contest Winner [31]	1.08	0.57	2.24	4.33
IR-Fusion (Ours)	**0.72**	**0.71**	6.98	**3.05**

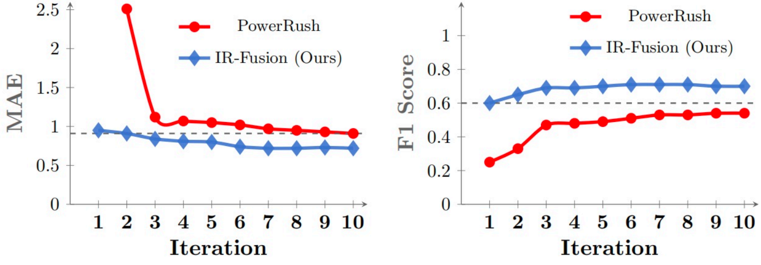


(a) Golden (b) MAUnet (c) IR-Fusion (Ours)

Fig. 6 Visualization of IR Drop distribution of a PG.

Experiment 2: Trade-off Study

- IR-Fusion surpasses PowerRush in all evaluated metrics.
- A key advantage of IR-Fusion is its ability to achieve the same MAE in just 2 iterations, while PowerRush requires 10 iterations to reach the same level.
- IR-Fusion consistently achieves a higher F1 score— a performance level PowerRush cannot reach at any iteration.
- Thanks to the fusion of numerical and ML methods, IR-Fusion achieves a better trade-off between accuracy and efficiency.



(a) MAE (b) F1 Score

Fig. 7 The analysis results of IR-Fusion and PowerRush [15].

Experiment 3: Ablation Study

- w./o. shows the results without a certain technique.
- Various techniques in our IR-Fusion contribute to the improvement of results significantly.

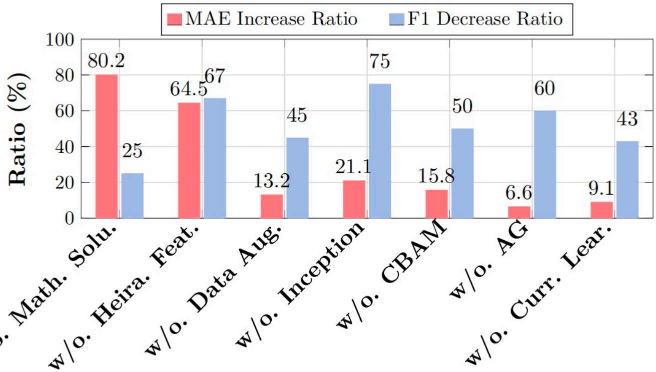


Fig. 8 Results of ablation study, showing the increase ratios in MAE (red) and decrease ratios in F1 score (blue).

Conclusion

Conclusion

- We propose IR-Fusion, an innovative fusion framework that incorporates numerical solutions to enhance ML methods for static IR drop analysis, providing an effective trade-off between accuracy and efficiency.
- Extensive experiments have demonstrated the **effectiveness and efficiency** of IR-Fusion.
- In the future, we hope to apply GNNs to capture PG's topology combing with the current image-based methods.

THANK YOU!