

Article

Fast Dynamic IR-Drop Prediction Using Machine Learning in Bulk FinFET Technologies

Pengcheng Huang *, Chiyuan Ma and Zhenyu Wu

Microelectronics and Microprocessor Institute, School of Computer, National University of Defense Technology, Changsha 410073, China; cyma@nudt.edu.cn (C.M.); zxdwzy@foxmail.com (Z.W.)

* Correspondence: huang.pc20@nudt.edu.cn

Abstract: IR-drop is a fundamental constraint by almost all integrated circuits (ICs) physical designs, and many iterations of timing engineer change order (ECO), IR-drop ECO, or other ECO are needed before design signoff. However, IR-drop analysis usually takes a long time and wastes so many resources. In this work, we develop a fast dynamic IR-drop predictor based on a machine learning technique, XGBoost, and the prediction method can be applied to vector-based and vectorless IR-drop analysis simultaneously. Correlation coefficient is often used to characterize the symmetry of prediction data and golden data, and our experiments show that the prediction correlation coefficient is more than 0.96 and the average error is no more than 1.3 mV for two industry designs, which are of 2.4 million and 3.7 million instances, respectively, and that the analysis is speeded up over 4.3 times compared with the IR-drop analysis by commercial tool, Redhawk.

Keywords: IR-drop (IRD); machine learning; engineer change order (ECO); XGBoost



Citation: Huang, P.; Ma, C.; Wu, Z.

Fast Dynamic IR-Drop Prediction
Using Machine Learning in Bulk
FinFET Technologies. *Symmetry* **2021**,
13, 1807. [https://doi.org/10.3390/
sym13101807](https://doi.org/10.3390/sym13101807)

Academic Editor: Christos Volos

Received: 23 August 2021

Accepted: 21 September 2021

Published: 28 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With technologies scaling down continuously, IR-drop, especially dynamic IR-drop, has brought about great concern for very large-scale integration (VLSI) physical designers and testers [1–4]. Lower voltage supply makes the signal slower [5] and IR-drop makes the cell delay larger [6], then decreasing the circuit performance or leading to timing failure or even increasing yield loss. Therefore, IR-drop signoff has become a significant step of physical design before tapping out. However, IR-drop analysis was often conducted by simulation-based commercial electronic design automation (EDA) tools, which were known to be accurate but very time-consuming. For a concrete industrial design, each iteration takes hours. Because no designer can hardly fix all IR-drop violations during one iteration, there needs to be dozens of ECO iterations to fix IR-drop violations, timing violations as well as design rule violations.

Figure 1 has shown the traditional IR-drop signoff flow for industrial design. The imported design data in one IR-drop analysis were very similar with that in the next IR-drop analysis because one ECO iteration only changes several to dozens of thousands of cells and nets, which is often less than 1% of the total design. Therefore, machine learning is suitable for predicting the IR-drop value of those ECO cells and then accelerates the flow runtime.

Until now, machine learning has been explored in IR-drop analysis. Linear regression is one of the simplest machine learning models. It was used to predict the IR-drop of each instance for at-speed scan test [6], for the power in scan-based testing is of great concern. The support vector machine (SVM) is another important machine learning model which was used to predict IR-drop [7,8]. The two kinds of algorithms were useful in some special designs and test fields, in which the IR-drop behaves in a linear relationship with the input features. Convolutional neural network (CNN) is a widespread machine learning model used in many research fields, especially image and voice recognition, and it was used for IR-drop prediction as well [9,10] because an IR-drop map can be

expressed as a figure. Regression tree is another widespread used machine learning model for many industrial application fields and it was also used for IR-drop estimation [10–13]. XGBoost [14], a gradient boosting regression tree, was proposed in 2016, which is proved to be behaving perfectly in most non-image application fields. In addition, XGBoost has achieved amazing results compared with other decision trees or even with some deep learning algorithms [10,15], and it has been used for IR-drop prediction as well [11–13].

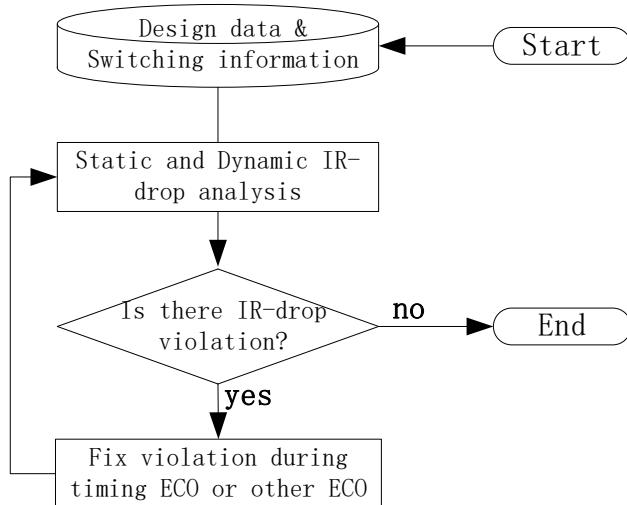


Figure 1. Traditional IR-drop signoff flow in physical design.

A major weakness shared by most of the previous works was that they focused on vector-based analysis but ignored vectorless IR-drop. On one hand, vector-based analysis depends on simulation patterns from value change dump (VCD) files, thus vector-based analysis is also called VCD-based analysis. However, while vectorless analysis does not need VCD files for dynamic IR-drop analysis, the physical designer is often unable to obtain accurate power simulation patterns in the early and middle stage of the design process, therefore vectorless analysis is more desirable for power analysis especially in the early stage of the design process. On another hand, the IR-drop from vectorless analysis and vector-based analysis is not the same, and vector-based analysis cannot cover all application scenarios, thus the physical designer needs to analyze vectorless and vector-based dynamic IR-drop before tapping out.

To overcome this shortcoming, in this paper we have proposed a machine learning method based on the XGBoost framework which can provide a simultaneous prediction of vectorless IR-drop and vector-based IR-drop. We put more emphasis on simultaneous prediction in our experiments, considering industrial demand and its higher difficulty and usability. The proposed method in this work addresses these challenges by its innovative pre-processed features and XGBoost algorithm. In addition, our experiment results also show that there is no need to update the machine learning model after each iteration of timing ECO or other ECO; it would be better to be updated for industrial designs after 10~30 thousands of ECO actions. Therefore, we can accelerate the design process with the proposed technique to cut down iteration time and we just run commercial IR-drop analysis tools when updating the machine learning model and signoff.

The rest of this paper is organized as follows. Section 2 shows the proposed technique. Section 3 presents the experiment setup. Section 4 shows experimental results based on two industry designs. Finally, Section 5 concludes this paper.

2. Proposed Technique

2.1. Proposed Flow

The proposed IR-drop signoff flow is shown in Figure 2. In the early stage of physical design, dynamic IR-drop analysis should be conducted using commercial EDA tools such

as Redhawk [16]. Then, based on the design data and power data from IR-drop analysis, some features should be extracted for each cell to train a machine learning model. Later, as the design is improved gradually, the IR-drop prediction can be carried out with the trained model, and then the designer can use the prediction data to fix IR-drop violations after each iteration of timing ECO or IR-drop ECO.

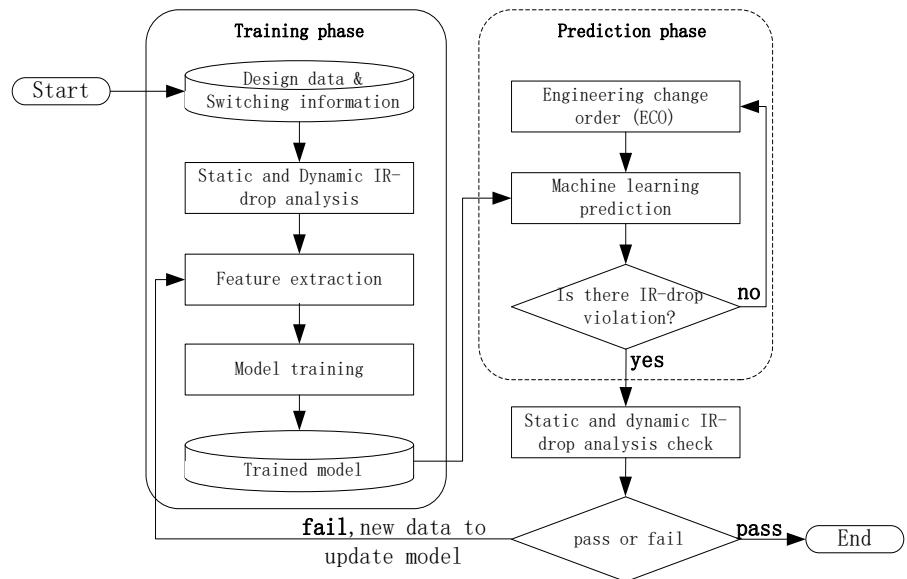


Figure 2. The proposed IR-drop signoff flow in physical design.

In the middle stage of physical design, the register transfer level (RTL) netlist and design floorplan are fixed, and then when the design data is also stable, the physical designer is recommended to run IR-drop analysis with commercial tool to update the trained model, leading to improved predicting accuracy in the later stage. Before signoff, if all predicted IR-drop violations by machine learning are fixed, the designer should perform IR-drop analysis with a commercial tool as well to check whether all real IR-drop violations are truly fixed. If there are true violations, the designer should update the trained model and repeat the aforementioned steps.

2.2. Nature Feature Extraction

Excessive IR-drop is caused by either larger resistance or larger current according to Ohm's Law, therefore an excellent design is of suitable power delivery network (PDN) and suitable local current density. PDN with uniform power grid means uniform resistance distribution in general. However, many designs need a non-uniform power grid, thus the resistance distribution is not uniform anymore and the pull-up resistance and pull-down resistance of each cell should be considered. The pull-up resistance and pull-down resistance are named R_{pu} and R_{pd} , respectively, and they can be extracted by a commercial tool such as Redhawk.

Current is a more critical factor related to IR-drop, but the current of each cell or each instance is not a constant value for the equivalent resistance of each cell, or each instance is not constant. The current of each cell is composed of static current and dynamic current; the static current is constant and is also called leakage current (I_l), while the dynamic current (I_{dyn}) varies as a signal switches. The scaled dynamic current (SI_{dyn}) is impacted by the signal toggle rate (Tog), and it is calculated by $I_{dyn} \times Tog$, and scaled total current (SI_A) is calculated by $SI_{dyn} + I_l$. SI_A and I_l are both extracted as input features. In addition, the peak current (I_p) has a significant effect on dynamic IR-drop so it must be extracted as a feature as well.

Since cell power consumption is proportional to cell current, cell power is taken into account as an input feature. There are three types of power consumptions, called leakage

power (P_l), internal power (P_i), and switching power (P_s), respectively. The leakage power P_l is consumed by unintended leakage, the internal power P_i is the power dissipated by internal parasitic capacitance of each cell, and the switching power P_s is the power dissipated by the output load capacitance of each cell. During dynamic IR-drop analysis, the internal power and switching power is scaled by toggle rate, and the toggle rate should be extracted as a feature as well. Since the toggle rate varies as VCD files vary, the scaled P_i (SP_i) and scaled P_s (SP_s) are calculated by $P_i \times Tog$ and $P_s \times Tog$, and the total scaled power (SP_A) is calculated by $SP_i + SP_s + P_l$. SP_i , SP_s , P_l , and SP_A are all extracted as input features.

The switching power is related with the output load capacitance and the transitions of all input signals. If a cell is moved to another location during ECO iteration, the drive load of the cell is changed but we cannot obtain an updated switching power based on static timing analysis (STA) data. Thus, the output load capacitance (C_L) should be extracted as an input feature. Besides, the number of input pins of each cell is not the same, and then the transitions of input signals are inconvenient as input features, and we do not consider it.

During each timing ECO or other ECO action, the location of some instance may be changed, and the cell type of some instance may be changed, thus some instance may be inserted or removed. Therefore, the x , y coordinate of a cell instance and the cell type of a cell instance should be extracted as input features. The cell type often includes three factors: the function such as AND or OR, the drive strength, and the voltage threshold (VT) type. Because most ECO actions only change the VT type and drive strength with the function keeping the same, the cell type can be split into three features: cell function (CF), drive strength (DS), and VT type (VTT). In total, all nature features extracted are shown in Table 1.

Table 1. Specifications of the extracted nature features from each instance itself.

Nature Features	Specifications
x, y	The x , y coordinate of an instance on layout design
CF	The code that represents the cell function of an instance
DS	The drive strength of an instance
VTT	The voltage threshold type of an instance
R_{pu}	The effective pull-up resistance of an instance
R_{pd}	The effective pull-down resistance of an instance
CL	The output load capacitance of an instance
Tog	The signal toggle rate of an instance
I_l	The leakage current of an instance
SI_A	The total of scaled dynamic current and leakage current of an instance
I_p	The peak current of an instance
P_l	The leakage power of an instance
SP_i	The scaled internal power of an instance
SP_s	The scaled switching power of an instance
SP_A	The scaled total power of an instance

2.3. Neighboring Feature Construction

The IR-drop of an instance is also impacted by its neighbor cells, as shown in Figure 3. The IR-drop of inst_A may be larger than that of inst_B though the total power of inst_A is less than inst_B for the cell density of inst_A is larger. Therefore, the power features of the neighboring cell should be considered. The cell density is not uniform everywhere, so we use 11×11 grids to construct the same number of neighboring features for each cell. Figure 3 illustrates smaller version (3×7) grids, where the coordinate of the center grid is the coordinate of the target cell. The SP_A of the grid is the sum of the product of the SP_A of each instance and its occupied area ratio related with this grid. For inst_A in Figure 3, it lies across two grids, 30% area of inst_A is in the left grid and 70% area of inst_A in the right grid, and 10% area of inst_B locates at the same grid with inst_A. Therefore, the SP_A of the left grid and the right grid is $0.3 \times SP_A(\text{inst_A})$ and $0.7 \times SP_A(\text{inst_A}) + 0.1 \times SP_A(\text{inst_B})$ respectively.

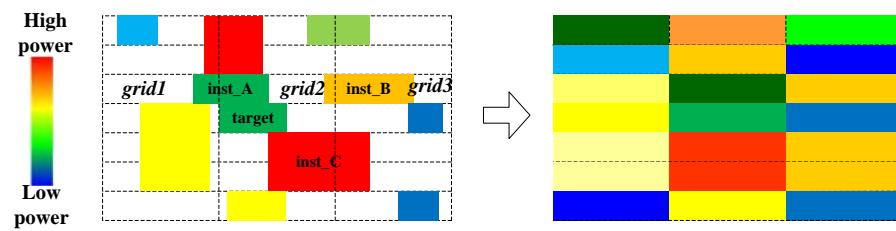


Figure 3. Power density map of SP_A and its transferred power grid map. The SP_A of the grid is the sum of the product of the SP_A of each instance and its occupied area ratio related with this grid. A total of 30% area of inst_A is in grid1, 10% area of inst_B is in grid2, so the SP_A of grid1, grid2, and grid3 is $0.3 \times SP_A(\text{inst_A})$, $0.7 \times SP_A(\text{inst_A}) + 0.1 \times SP_A(\text{inst_B})$, $0.9 \times SP_A(\text{inst_B})$, respectively.

In addition, there exist multi-row standard cells in commercial bulk FinFET technology, such as inst_C shown in Figure 3. The power calculation is similar to those single-row cells. The height of each grid is equal to the height of a standard cell row and the width of each grid is several or decades times of one site's width. In this work, the height of each grid is 0.3 μm , and the width of each grid is 1.28 μm , which is 20 times of the one site's width. The selection of grid numbers depends on experience, and it can be varied with the change of technology.

Toggle rate Tog is also used to build neighboring features to improve prediction accuracy and its calculation is the same as that for SP_A .

The density of the scaled total current SI_A and the peak current I_p should also be considered, and they are extracted using 11×11 grids as well. However, the relationship between IRD with current is not the same as that between IRD with power assumption. As shown in Figure 4, the IRD of inst1 is proportional to the sum of $I1_{up}$, $I1_{down}$, $I2_{up}$, and $I3_{down}$. For simplicity, we suppose that the effect of $I1_{up}$ and $I1_{down}$ is the same and that they are the half of the current of inst1. For I_p and SI_A of inst1 in Figure 4, we can achieve I_p' and SI_A' of inst1, which are $I_p1 + (I_p2 + I_p3)/2$, and $SI_A1 + (SI_A2 + SI_A3)/2$, respectively. For neighboring features about peak current and total current, we also use 11×11 grids, and the same calculation method is adopted, but the parameters I_p and SI_A are replaced by I_p' and SI_A' .

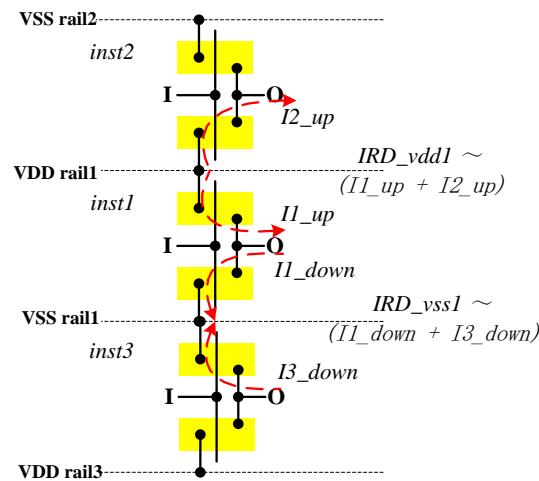


Figure 4. The relationship between IRD of an instance and the current of related instances. IRD of inst1 is composed of IRD_{vdd1} and IRD_{vss1} . IRD_{vdd1} is related with the pull-up current of inst1 and inst2, i.e., $I1_{up}$ and $I2_{up}$. IRD_{vss1} is related with the pull-down current of inst1 and inst2, i.e., $I1_{down}$ and $I3_{down}$. Therefore, the IRD of inst1 can be supposed to be proportional to $(I1_{up} + I1_{down} + I2_{up} + I3_{down})$.

Totally, there are 484 ($11 \times 11 \times 4$) neighboring features extracted as well as 16 nature features. Totally, there are 500 features in total for each cell.

2.4. Design Matrix Construction

Many designs have large amounts of instances, for example, several million or dozens of millions. If all instances are considered simultaneously, the design matrix is too large. In order to mitigate the training scale, many researchers are inclined to partition the floorplan and then select parts of IRD critical partitions to construct a training model, such as in [10–13]. This method can really save training time, but the cell IRD in other unselected partitions is not completely considered. Besides, during hold ECO, hold buffers are inserted almost everywhere and there are ECO actions in each partition. Then the partition method will overlook part of IRD prediction of ECO cells or other cells neighboring the ECO cells. Therefore, in this work, we divide the floorplan into several partitions so that there exists 0.4–1 million instances in each partition. For each partition, we combine the top selection with random selection to construct the design matrix.

For vectorless and VCD-based dynamic IRD prediction, the top selection means selecting 120,000 instances of which the IR-drop is the worst in the partition and the random selection means selecting another 180,000 instances randomly in the partition. By this method, we use 40% samples to characterize the worst IR-drop cells and 60% samples to characterize other cells in each partition, and then the IRD value from zero to the worst are all considered. In total, 300,000 samples are selected to construct a design matrix for each partition, 40% samples are from the top selection and 60% samples are from a random selection. The selection number is determined based on experience. Besides, for those blocks or partitions with less than 300,000 instances, all instances can be selected into the design matrix.

2.5. Training Model

We use XGBoost as the machine learning models. A given dataset is of form $D = (x_i, y_i)$, where $|D| = N$, $x_i \in R^m$, $y_i \in R$, N is the total number of cells, m is the feature dimension, x_i is just the input features of the i th instance, and y_i is just the golden IRD value of i th instance provided by commercial tools such as Redhawk. The XGBoost ensembles K additive functions to achieve the predicted IRD value by Formula (1).

$$\tilde{y}_i = \Phi(x) = \sum_{k=1}^K f_k(x_i), f_k \in F \quad (1)$$

where F is the space of regression trees and f_k is k th tree function. The goal is to minimize the following objective function shown in Formula (2).

$$L(\Phi) = \sum_i l(y_i, \tilde{y}_i) + \sum_k \Omega(f_k) \quad (2)$$

where each $l(y_i, \tilde{y}_i)$ is a differentiable convex loss function that characterizes the difference between the predicted IRD \tilde{y}_i and the golden IRD y_i . In our application, it is defined by mean square error (MSE). $\Omega(f_k)$ is a regularization term that penalizes the complexity of the tree structure and prevents overfitting.

3. Experiments Setup

Two industry design blocks, DesignA and DesignB, have been evaluated by our proposed method. The voltage supply for both designs is 0.75 V, and the profiles of these two circuits are shown in Table 2. The second column shows the dynamic analysis mode where one VCD is available for DesignA while two VCDs are available for DesignB. For both designs, the golden IRD values are provided by Redhawk, the mean IRD and the max IRD values are shown in the fourth and fifth column, respectively. The number of IRD violations means an IRD value larger than 8% of the voltage supply in dynamic mode, which is shown in the sixth column.

Table 2. The profiles of evaluated circuit designs, i.e., DesignA and DesignB.

Circuit Design	DesignA		DesignB		
Power Analysis Mode	Vectorless	VCD	Vectorless	VCD1	VCD2
Number of cell instance	2.4 million	2.4 million	3.7 million	3.7 million	3.7 million
Mean IRD (mV)	15.31	2.19	16.84	3.54	8.42
Max IRD (mV)	55.7	165.2	73.5	90.9	91.4
Number of IRD violations	0	576	9	28	173

We use correlation coefficient (CC), Mean Absolute Error (MAE), Max Error (MaxE), Root Mean Square Error (RMSE), and Normalized Root Mean Square Error (NRMSE) to evaluate the performance of our method and which are defined in following equations. CC is defined in Equation (3) and it is a number between 0 and 1, where 1 means completely correlated. MAE and MaxE are defined in Equations (4) and (5), RMSE and NRMSE are defined in Equations (6) and (7).

$$CC = \frac{\sum_{i=1}^N (y_i - \text{mean}(y))(\tilde{y}_i - \text{mean}(\tilde{y}))}{\sqrt{\sum_{i=1}^N (y_i - \text{mean}(y))^2 \sum_{i=1}^N (\tilde{y}_i - \text{mean}(\tilde{y}))^2}} \quad (3)$$

$$MAE = \frac{\sum_{i=1}^N ||y_i - \tilde{y}_i||}{N} \quad (4)$$

$$MaxE = \max_{i \in [1, N]} (y_i - \tilde{y}_i) \quad (5)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \tilde{y}_i)^2}{N}} \quad (6)$$

$$NRMSE = \frac{RMSE}{\text{mean}(y)} \times 100\% \quad (7)$$

4. Results

4.1. IR-Drop Prediction before ECO

The design data can be from any design database in which the routing is finished. In this work, the data are from the database where placement and routing (PR) are finished and no ECO action is carried out. For DesignA and DesignB, there are four partitions shown in Figure 5. The training data are generated by top selection and random selection described in the context, the training data numbers are 300 k, which is about 1/4–2/3 of total cell instances in each partition in DesignA and DesignB, respectively. We also select the total instances in each partition to construct one group of testing data to validate whether our model can successfully predict all IRD values, and the results are shown in Tables 3 and 4.

As shown in Table 3, the MAE is less than 0.42 mV and the CC is larger than 0.982 for VCD-based dynamic mode in each partition for DesignA. The MAE is less than 0.847 mV and the CC is larger than 0.974 for vectorless dynamic mode in each partition.

As shown in Table 4, the MAE is less than 1116 mV and the CC is larger than 0.961 for VCD1-based dynamic mode in each partition for DesignB. The MAE is less than 1.26 mV and the CC is larger than 0.960 for vectorless dynamic mode in each partition.

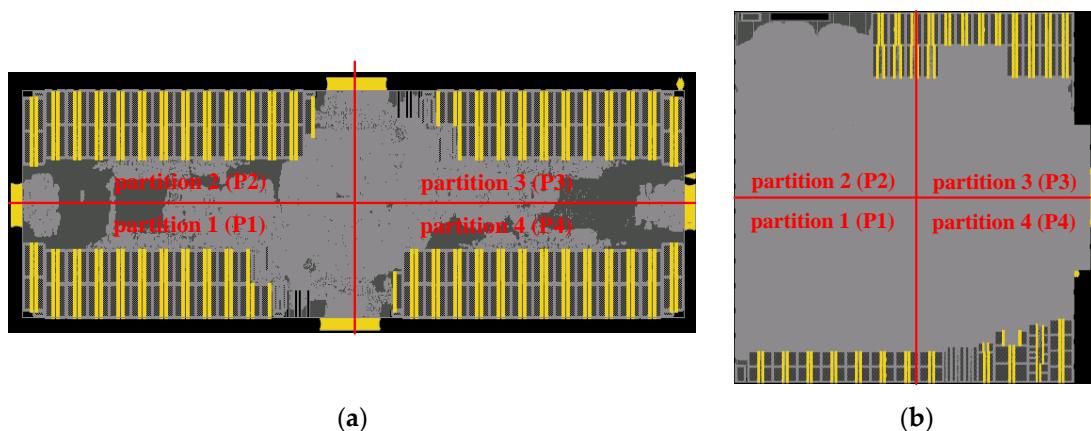


Figure 5. The partition for DesignA and DesignB: (a) there are 4 partitions in DesignA, which is recoded as P1–P4; (b) there are 4 partitions in DesignB, which is recoded as P1–P4 as well.

Table 3. The prediction results of DesignA before ECO.

Partitions	Power Analysis Mode	Inst Numbers	Mean IRD (mV)	CC	MAE (mV)	MaxE (mV)	RMSE
P1	vectorless	545,772	15.97	0.976	0.845	15.49	1.14
	VCD	545,772	1.69	0.982	0.267	7.10	0.38
P2	vectorless	444,511	14.69	0.982	0.784	16.03	1.07
	VCD	444,511	1.81	0.986	0.291	8.95	0.43
P3	vectorless	590,385	16.03	0.974	0.847	16.55	1.14
	VCD	590,385	2.74	0.988	0.420	14.14	0.59
P4	vectorless	460,383	14.22	0.979	0.807	14.89	1.11
	VCD	360,383	2.43	0.992	0.333	10.88	0.49

Table 4. The prediction results of DesignB before ECO.

Partitions	Power Analysis Mode	Inst Numbers	Mean IRD (mV)	CC	MAE (mV)	MaxE (mV)	RMSE
P1	vectorless	1,177,525	16.37	0.960	1260	21.53	1.67
	VCD1	1,177,525	1.70	0.976	0.312	14.29	0.48
	VCD2	1,177,525	11.38	0.972	1116	21.44	1.57
P2	vectorless	947,562	14.13	0.960	1087	21.92	1.47
	VCD1	947,562	2.01	0.987	0.276	18.09	0.45
	VCD2	947,562	4.13	0.961	0.777	22.17	1.17
P3	vectorless	782,721	16.81	0.963	1176	23.44	1.58
	VCD1	782,721	6.87	0.978	0.766	23.12	1.12
	VCD2	782,721	9.63	0.980	1077	26.09	1.58
P4	vectorless	807,029	20.42	0.981	1092	19.81	1.47
	VCD1	807,029	4.87	0.974	0.652	23.79	0.96
	VCD2	807,029	7.84	0.965	0.923	25.60	1.31

Figure 6 shows the correlation between the predicted IRD and the golden IRD for DesignA. Figure 6a–d is the result for VCD-based dynamic mode while Figure 6e–h is for vectorless dynamic mode. The results show high correlation of our prediction in both dynamic modes. For vectorless dynamic mode, as shown in Figure 6e–h, the correlation of the prediction for low IRD cells is smaller than that for high IRD cells. This is due to the selection for low IRD cells being far less than 70%. However, the designer does focus on the prediction of high IRD and IRD violations. The limited loss of prediction accuracy of low IRD has no effect on the prediction of high IRD and we would rather accelerate the training process by this method.

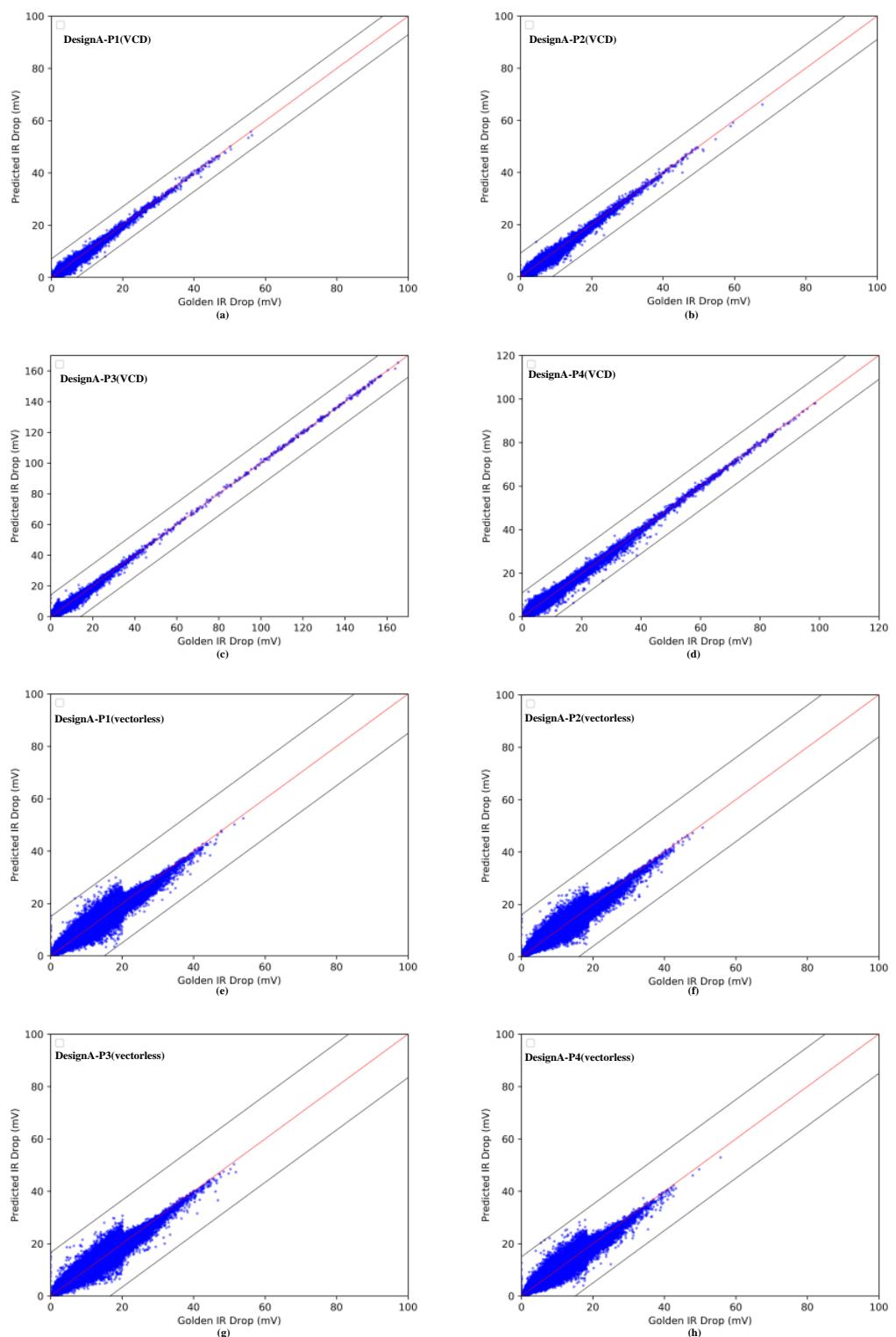


Figure 6. The predicted IR-drop versus golden IR-drop for DesignA before ECO, (a) for Partition 1 with VCD-based dynamic mode, (b) for Partition 2 with VCD-based dynamic mode, (c) for Partition 3 with VCD-based dynamic mode, (d) for Partition 4 with VCD-based dynamic mode, (e) for Partition 1 with vectorless dynamic mode, (f) for Partition 2 with vectorless dynamic mode, (g) for Partition 3 with vectorless dynamic mode, (h) for Partition 4 with vectorless dynamic mode.

Because the cell numbers in each partition of DesignB are much larger than that of DesignA, the CC and MAE are often larger than DesignA in each partition, that is to say,

the prediction accuracy of DesignB is decreased slightly due to the larger design volume. However, as shown in Figure 7, the correlation of high IRD values is high though the correlation of low IRD values is decreased due to the larger design volume, which means the proposed method can be competent at the prediction accuracy of high IRD values and IRD violations.

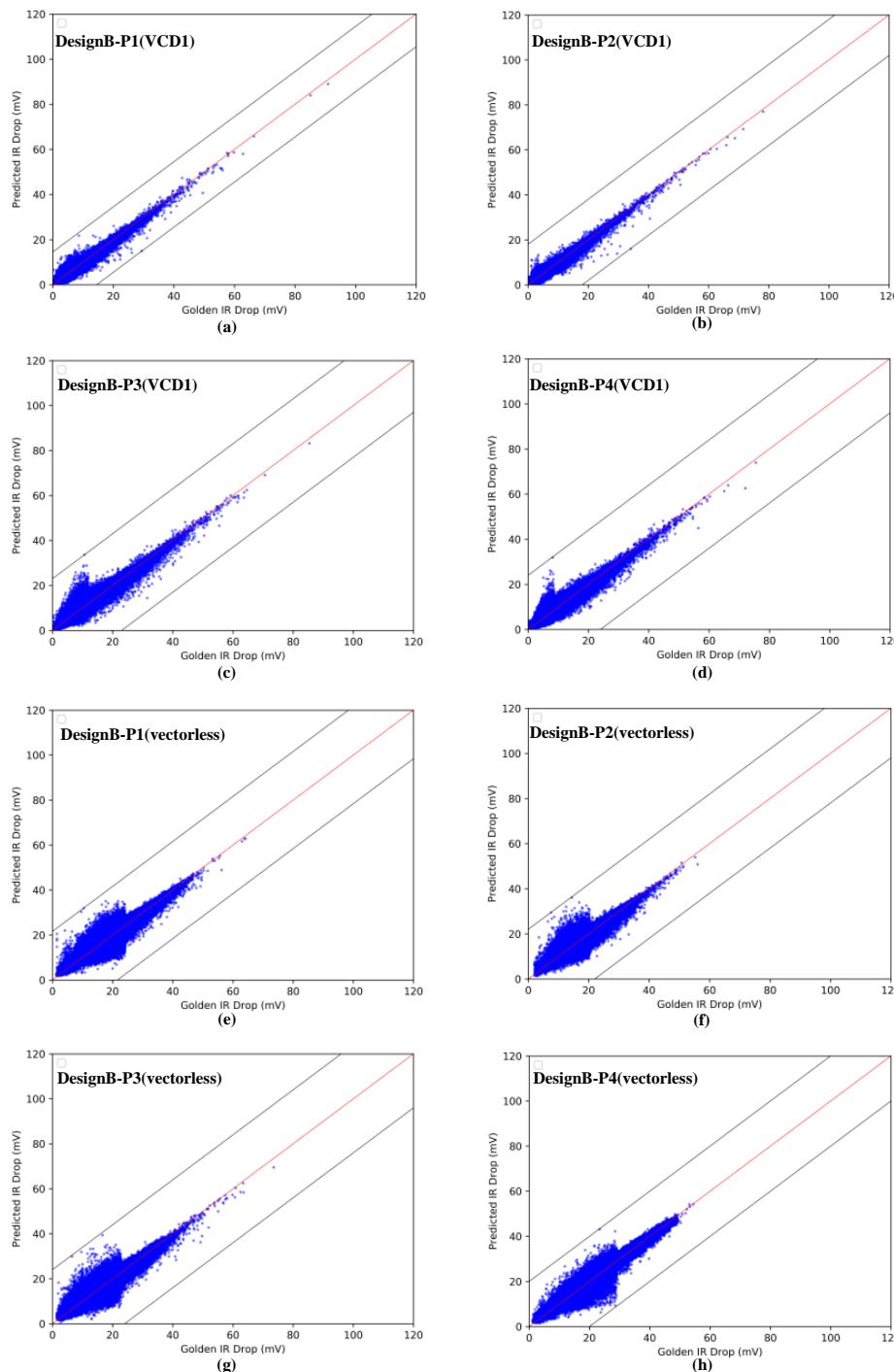


Figure 7. The predicted IR-drop versus golden IR-drop for DesignB before ECO, (a) for Partition 1 with VCD-based dynamic mode, (b) for Partition 2 with VCD-based dynamic mode, (c) for Partition 3 with VCD-based dynamic mode, (d) for Partition 4 with VCD-based dynamic mode, (e) for Partition 1 with vectorless dynamic mode, (f) for Partition 2 with vectorless dynamic mode, (g) for Partition 3 with vectorless dynamic mode, (h) for Partition 4 with vectorless dynamic mode.

Figure 8 shows the error distribution of predicted IRD for DesignA and DesignB. The Y-axis is the number of cell instances, and the X-axis of the figure is golden IRD minus predicted IRD. The left part of each figure in Figure 8 shows the overpredicted cell instances while the right part of each figure shows the underpredicted cell instances. Most of the prediction (>99%) errors are smaller than ± 5 mV for DesignA and DesignB for all dynamic modes.

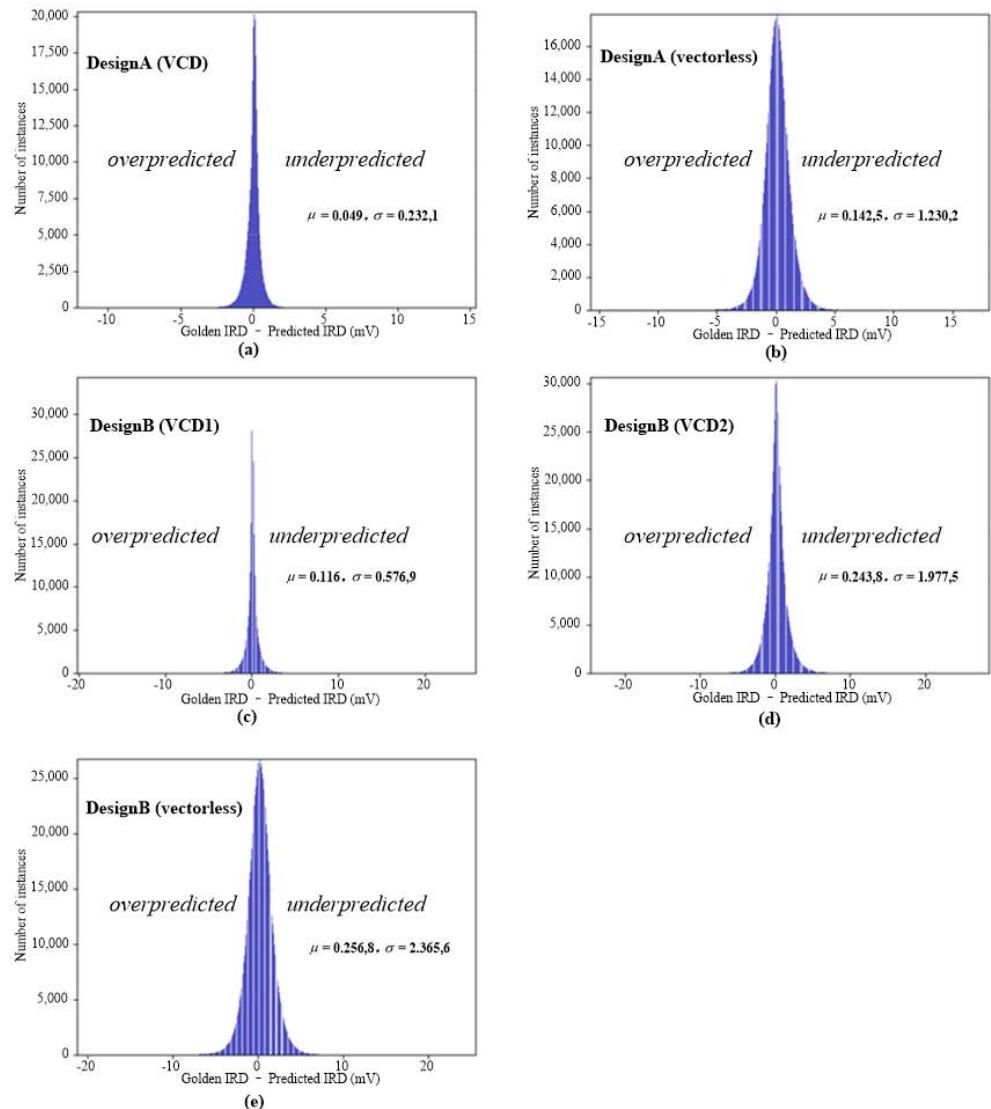


Figure 8. The error distribution of the prediction before ECO, (a) for DesignA with VCD-based dynamic mode, (b) for DesignA with vectorless dynamic mode, (c) for DesignB with dynamic mode with VCD1, (d) for DesignB with dynamic mode with VCD2, (e) DesignB with vectorless dynamic mode.

4.2. IR-Drop Prediction after ECO

Several iterations of ECO actions have been completed on the design database of DesignA and DesignB, respectively. The ECO actions include sizing-up or sizing-down cells, inserting buffer, removing buffer, VT-swapping and moving cells, aiming to fix timing, and fixing some IR-drop violations. Totally, there are 10,057 and 91,916 ECO actions for DesignA and DesignB, respectively. We used the trained XGBoost model to predict the IR-drop of the cell instances in ECO database and we obtained the golden IR-drop of all cells from Redhawk.

The comparison of the predicted IRD with the golden IRD is shown in Tables 5 and 6. As shown in Table 5, the CC is larger than 0.901 and 0.971 for VCD-based and vectorless mode, respectively, and the MAE is less than 1189 mV and 0.888 mV for VCD-based and vectorless mode, respectively. As shown in Table 6, the CC is larger than 0.915 and 0.882 for VCD-based and vectorless mode, respectively, and the MAE is less than 1955 mV and 2370 mV for VCD-based and vectorless mode, respectively. Compared with the predicted results shown in Tables 3 and 4, the CC is decreased slightly, MAE is increased, and MaxE is also increased. That is to say, the prediction accuracy is decreased slightly due to dozens of thousands of ECO actions.

Table 5. The prediction results of DesignA after ECO.

Partitions	Power Analysis Mode	Inst Numbers	Mean IRD (mV)	CC	MAE (mV)	MaxE (mV)	RMSE
P1	vectorless	545,795	16.03	0.973	0.888	16.63	1.19
	VCD	545,795	1.74	0.970	0.320	20.53	0.52
P2	vectorless	444,543	14.75	0.980	0.835	18.86	1.14
	VCD	444,543	1.80	0.979	0.324	13.3	0.53
P3	vectorless	590,410	16.10	0.971	0.882	20.97	1.19
	VCD	590,410	3.98	0.904	1189	22.45	1.88
P4	vectorless	460,400	14.12	0.975	0.887	16.30	1.20
	VCD	460,400	2.08	0.901	0.882	20.02	0.98

Table 6. The prediction results of DesignB after ECO.

Partitions	Power Analysis Mode	Inst Numbers	Mean IRD (mV)	CC	MAE (mV)	MaxE (mV)	RMSE
P1	vectorless	1,185,523	16.25	0.882	2.347	32.91	2.99
	VCD1	1,185,523	1.69	0.962	0.379	31.04	0.66
	VCD2	1,185,523	11.80	0.915	1955	35.96	2.88
P2	vectorless	959,577	14.15	0.908	2010	28.90	2.66
	VCD1	959,577	1.79	0.964	0.404	30.57	0.78
	VCD2	959,577	4.46	0.945	0.887	34.47	1.36
P3	vectorless	812,582	16.90	0.910	2370	36.53	3.11
	VCD1	812,582	7.12	0.963	1091	37.70	1.68
	VCD2	812,582	11.01	0.970	1700	39.91	3.08
P4	vectorless	829,071	20.63	0.929	2256	36.19	2.95
	VCD1	829,071	5.07	0.935	1049	30.80	1.73
	VCD2	829,071	5.48	0.921	1485	39.50	1.49

Figure 9 shows the difference between the predicted IRD with the golden IRD after ECO. For DesignA, after 10,057 ECO actions, the difference between the predicted high IRD and the golden high IRD is increased, especially in VCD-based dynamic mode. For DesignB, after more than 90,000 ECO actions, the difference between the predicted high IRD and the golden high IRD is increased as well, especially in dynamic analysis with VCD2. That is to say, the prediction accuracy of high IRD values and IRD violations is decreased after ECO.

However, as shown in Tables 5 and 6, the CC for most cases is still larger than 0.9 and the MAE is less than 1.2 mV and 2.4 mV for DesignA and DesignB, respectively. In addition, the MaxE is less than 21 mV and 40 mV for DesignA and DesignB, respectively, which means the prediction accuracy is still high.

In order to illustrate the prediction effect more clearly, the IRD maps are shown in Figures 10 and 11. Figure 10a is the predicted IRD map of DesignA in VCD-based dynamic mode, Figure 10b is the corresponding golden IRD map, and they are nearly the same. Figure 10c,d is the predicted IRD map and the golden IRD map of DesignA in vectorless dynamic mode, and they are nearly the same as well. Figure 11a,b is the predicted IRD map and the golden IRD map of DesignB in VCD1-based dynamic mode, which are nearly the same. Figure 11c,d is the predicted IRD map and golden IRD map of DesignB in

VCD2-based dynamic mode, which are also nearly the same. Figure 11e,f is the predicted IRD map and golden IRD map of DesignB in vectorless dynamic mode, which are nearly the same as well. Therefore, the prediction by the proposed model is also successful after dozens of thousands of ECO actions.

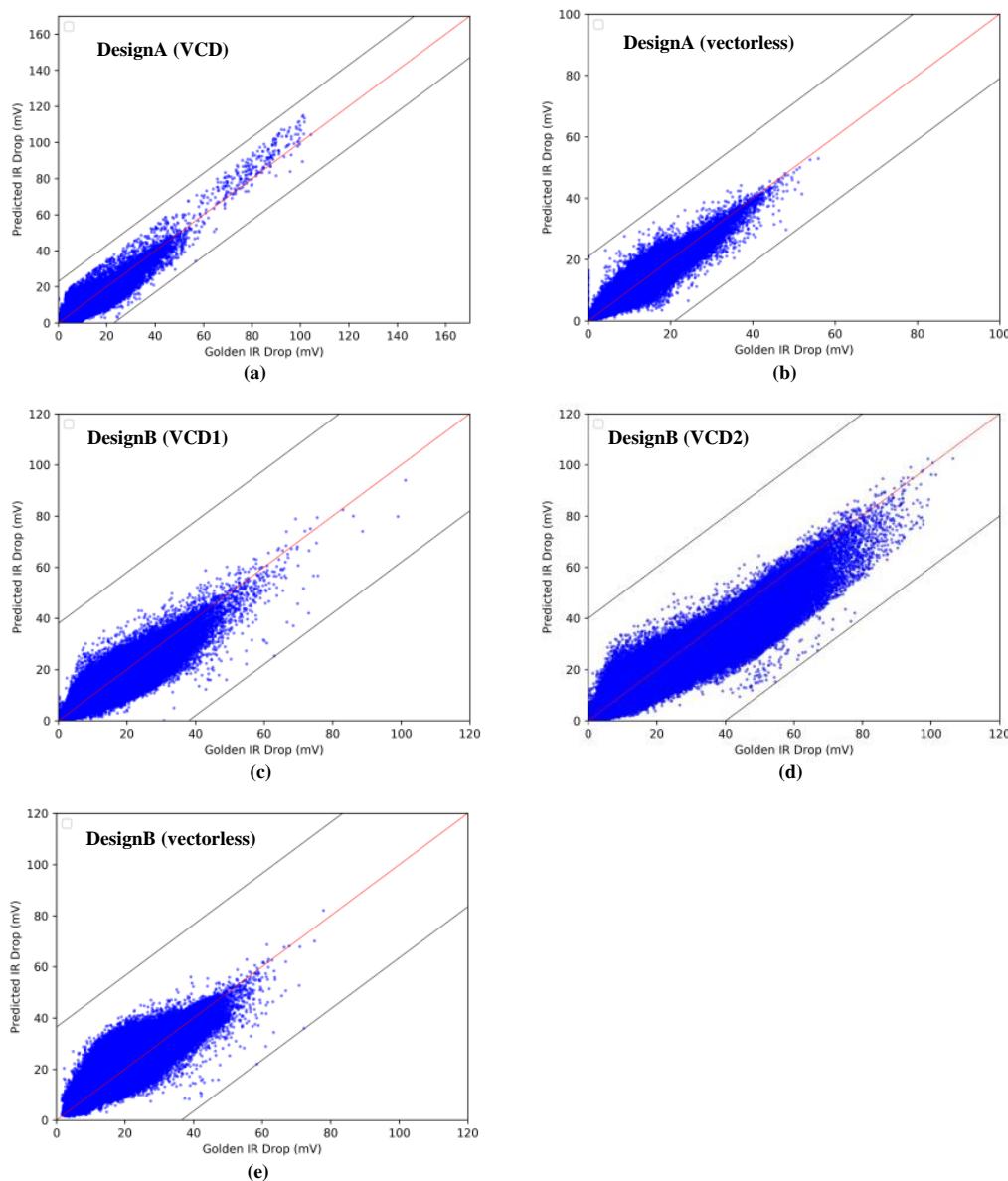


Figure 9. The predicted IR-drop versus golden IR-drop after ECO, (a) for DesignA with VCD-based dynamic mode, (b) for DesignA with vectorless dynamic mode, (c) for DesignB with dynamic mode with VCD1, (d) for DesignB with dynamic mode with VCD2, (e) for DesignB with vectorless dynamic mode.

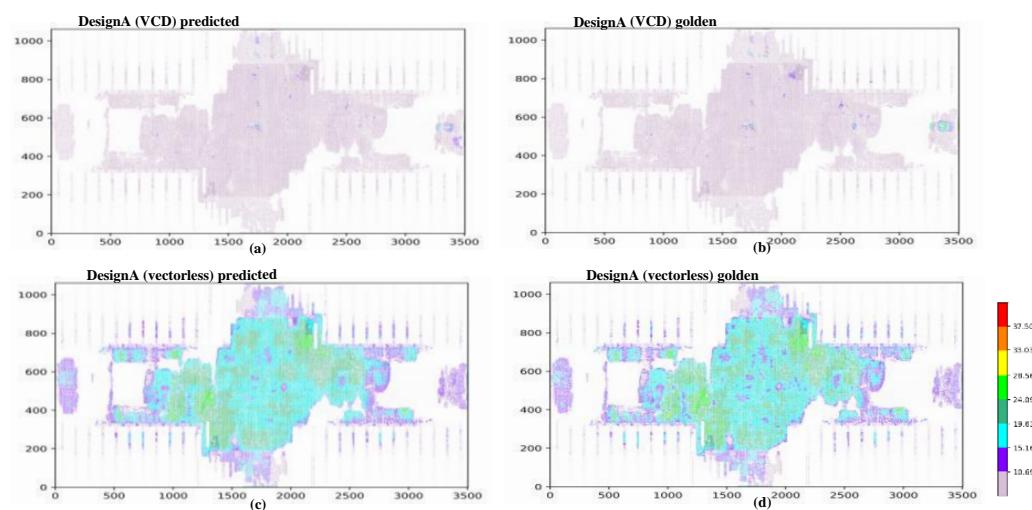


Figure 10. (a) The predicted IR-drop map for DesignA after ECO with VCD-based dynamic mode, (b) the golden IR-drop map for DesignA after ECO with VCD-based dynamic mode, (c) the predicted IR-drop map for DesignA after ECO with vectorless dynamic mode, (d) the golden IR-drop map for DesignA after ECO with vectorless dynamic mode.

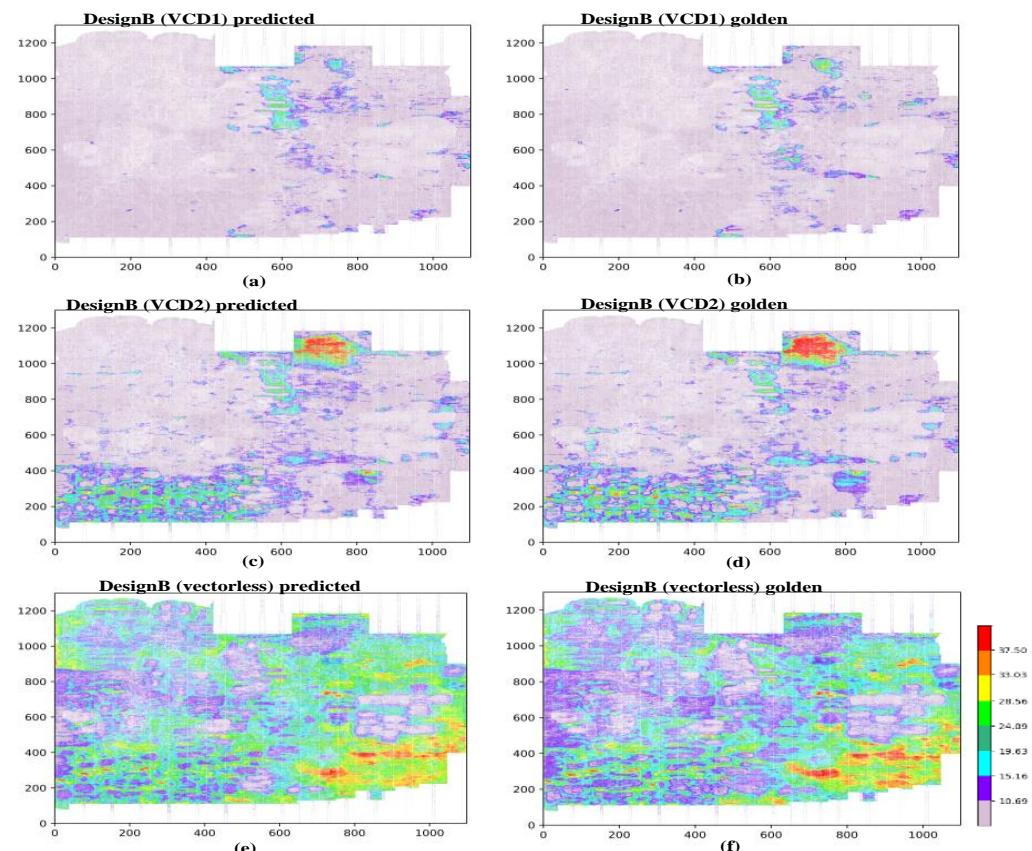


Figure 11. (a) The predicted IR-drop map for DesignB after ECO with dynamic mode with VCD1, (b) the golden IR-drop map for DesignB after ECO with dynamic mode with VCD1, (c) the predicted IR-drop map for DesignB after ECO with dynamic mode with VCD2, (d) the golden IR-drop map for DesignB after ECO with dynamic mode with VCD2, (e) the predicted IR-drop map for DesignB after ECO with vectorless dynamic mode, (f) the golden IR-drop map for DesignB after ECO with vectorless dynamic mode.

In order to validate whether the true IR-drop violations are predicted correctly furthermore, two more evaluation metrics, Precision and Recall, are introduced. Predicted IR pass (PP) means those cells of which the predicted IRD is less than 60 mV ($=8\% \times 750$ mV), Predicted IR Fail (PF) means those cells of which the predicted IRD is larger than 60 mV. Golden IR pass (GP) means those cells who's golden IRD is less than 60 mV while Golden IR Fail (GF) means those cells who's golden IRD is larger than 60 mV. Then, Precision and Recall are defined as follows.

$$\left\{ \begin{array}{l} TruePositive (TP) = PF \cap GF \\ TrueNegative (TN) = PP \cap GP \\ FalseNegative (FN) = PF \cap GP \\ FalsePositive (FP) = PP \cap GF \end{array} \right\} \quad (8)$$

$$\left\{ \begin{array}{l} Precision = \frac{|TP|}{|TP| + |FP|} \\ Recall = \frac{|TP|}{|TP| + |FN|} \end{array} \right\} \quad (9)$$

For DesignA and DesignB, the Precision and Recall is shown in Table 7. For DesignA, no IRD violation exists in vectorless dynamic mode, and the Precision and Recall is 0.95 and 0.98, respectively, in VCD-based dynamic mode, which means that 95% of the predicted IRD violations are really IRD violations, and that 98% of IRD violations are successfully identified. For DesignB, the Precision is larger than 0.794 for three dynamic modes, which means more than 79.4% of the predicted IRD violations are really IRD violations, and the Recall is over 0.382, which means more than 38.2% of IRD violations can be identified successfully. The Recall of DesignB is small because the ECO actions are more than 90,000, and then the training model should be updated based on the ECO database for DesignB to promote Recall for next ECO actions. From this view, maybe after 10,000 to 20,000 ECO actions, the training model should be updated.

Table 7. The Precision and Recall of two designs after ECO.

Circuit Design	MAE (mV)	MaxE (mV)	CC	RMSE	TP	TN	FP	FN	Precision	Recall
DesignA (vectorless)	0.875	20.97	0.975	1.181	0	2041148	0	0	NAN	NAN
DesignA (VCD)	0.643	22.45	0.925	1.209	246	2040884	13	5	0.950	0.980
DesignB (vectorless)	2.247	36.53	0.916	2.925	13	3786717	2	21	0.867	0.382
DesignB (VCD1)	0.739	37.70	0.954	1.410	27	3786678	7	41	0.794	0.397
DesignB (VCD2)	1.526	39.91	0.952	2.488	2137	3781511	143	2962	0.937	0.419

4.3. Runtime Comparison

The experiments are performed on an Intel Xeon Processor W-2245, and the runtime of the work is shown in Table 8. For DesignA, we spend 21–24 min to construct the training matrix, and the training time is about 30–31 min, thus the total training time is about 52~54 min. All instances in the design are taken into account during the predicting matrix construction so that this process cost over 1 h. The predicting time is less than 1 min. If we use Redhawk to run dynamic IRD analysis, we need about 5.5 h and 6.6 h for DesignA and DesignB, respectively. After one or several iterations of ECO, we should reconstruct the predicting matrix, then obtain the predicted IRD values of all instances with the training model, which cost about 1.18 h and 1.5 h, respectively. Therefore, the dynamic IRD analysis is speeded up over 4.3 times for both designs.

Table 8. The runtime comparison of two designs.

Circuit Design	DesignA (Vectorless)	DesignA (VCD)	DesignB (Vectorless)	DesignB (VCD1)	DesignB (VCD2)
Constructing training matrix	23 min 22 s	21 min 49 s	24 min 7 s	22 min 31 s	22 min 43 s
Training	30 min 11 s	30 min 7 s	30 min 12 s	30 min 8 s	30 min 9 s
Constructing predicting matrix	1 h 11 min	1 h 9 min	1 h 36 min	1 h 27 min	1 h 27 min
Predicting	31 s	30 s	39 s	38 s	38 s
IRD analysis by Redhawk	5 h 41 min	5 h 35 min	6 h 57 min	6 h 34 min	6 h 44 min

5. Conclusions

A machine-learning-based dynamic IRD predictor was proposed to speed up IRD analysis for ECO over 4.3 times. We proposed the construction approach of a design matrix based on partition, top selection, and random selection, which is the trade-off of prediction accuracy and training time. The extracted features take geometry size, power, and physical information of instances into account. The experiments show that the prediction CC is more than 0.96 and MAE is less than 1.3 mV on the 2.4 million-cells and 3.7 million-cells industry designs in all dynamic modes. Precision and Recall for IRD violations after over 10,000 ECO actions in DesignA are 0.95 and 0.98, and Precision for IRD violations after over 90,000 ECO actions in DesignB is still more than 0.79. We think it is suitable for the designer to update the machine learning model after 10~30 thousand ECO actions, and our directions for ongoing work include: (i) improving the speed of our method, (ii) improving the accuracy of the prediction by optimizing feature extraction, and (iii) extracting feature for power-mesh, allowing the machine learning model to be more suitable for non-uniform PDN.

Author Contributions: Conceptualization, P.H. and C.M.; methodology, P.H.; software, P.H.; validation, C.M. and Z.W.; investigation, P.H.; resources, P.H.; data curation, P.H. and Z.W.; writing—original draft preparation, P.H. and Z.W.; writing—review and editing, P.H.; visualization, P.H. and C.M.; funding acquisition, P.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Science Foundation of China, grant number 61704192.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to their storage in private networks.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Shepard, K.L.; Narayanan, V. Noise in deep submicron digital design. In Proceedings of the International Conference on Computer Aided Design, San Jose, CA, USA, 9–13 November 1997; pp. 524–531.
- Tehrani Poor, M.; Butler, K.M. Power supply noise: A survey on effects and research. *IEEE Des. Test Comput.* **2010**, *2*, 51–67. [[CrossRef](#)]
- Nithin, S.K.; Shanmugam, G.; Chandrasekar, S. Dynamic voltage (IR) drop analysis and design closure: Issues and challenges. In Proceedings of the IEEE International Symposium on Quality Electronic Design, San Jose, CA, USA, 15 April 2010; pp. 611–617.
- Holst, S.; Schneider, E.; Wen, X.; Kajihara, S.; Kochte, M.A. Timing-Accurate Estimation of IR-Drop Impact on Logic- and Clock-Paths during At-Speed Scan Test. In Proceedings of the IEEE Asian Test Symposium, Hiroshima, Japan, 21–24 November 2016; pp. 19–24.
- Pengcheng, H.; Shuming, C.; Jianjun, C. Simulation study of N-hit SET variation in differential cascade voltage switching logical circuits. *Sci. China Inf. Sci.* **2015**, *58*, 022401.
- Yamato, Y.; Yoneda, T.; Hatayama, K.; Inoue, M. A fast and accurate per-cell dynamic IR-drop estimation method for at-speed scan test pattern validation. In Proceedings of the IEEE International Test Conference, Anaheim, CA, USA, 5–8 November 2012; pp. 1–8.
- Chapelle, O.; Haffner, P.; Vapnik, V. Support vector machines for histogram-based image classification. *IEEE Trans. Neural Netw.* **1999**, *5*, 1055–1064. [[CrossRef](#)] [[PubMed](#)]
- Ye, F.; Firouzi, F.; Yang, Y.; Chakrabarty, K.; Tahoori, M.B. On-chip droop-induced circuit delay prediction based on support-vector machines. *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.* **2016**, *4*, 665–678. [[CrossRef](#)]
- Xie, Z.; Ren, H.; Khailany, B.; Sheng, Y.; Santosh, S.; Hu, J.; Chen, Y. PowerNet: Transferable dynamic IR drop estimation via maximum convolutional neural network. In Proceedings of the Asia and South Pacific Design Automation Conference, Beijing, China, 13–16 January 2020; pp. 13–18.
- Fang, Y.; Lin, H.; Su, M.; Li, C.; Fang, E.J. Machine-learning-based dynamic IR drop prediction for ECO. In Proceedings of the IEEE/ACM International Conference on Computer Aided Design, San Diego, CA, USA, 5–8 November 2018; pp. 1–7.
- Ho, C.; Kahng, A. IncPIRD: Fast learning-based prediction of incremental IR drop. In Proceedings of the IEEE/ACM International Conference on Computer Aided Design, Denver, CO, USA, 4–7 November 2019; pp. 1–8.

12. Lin, H.; Fang, Y.; Liu, S.; Chen, J.X.; Fang, J.W. Automatic IR-drop ECO using machine learning. In Proceedings of the IEEE International Test Conference in Asia, Taipei, Taiwan, 23–25 September 2020; pp. 7–12.
13. Pao, C.; Su, A.; Lee, Y. XGBIR: An XGBoost-based IR drop predictor for power delivery network. In Proceedings of the Design, Automation & Test in Europe Conference & Exhibition (DATE), Grenoble, France, 9–13 March 2020; pp. 1307–1310.
14. Chen, T.; Guestrin, G. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
15. Pu, Q.; Li, Y.; Zhang, H.; Yao, H.; Zhang, B.; Hou, B.; Li, L.; Zhao, Y.; Zhao, L. Screen efficiency comparisons of decision tree and neural network algorithms in machine learning assisted drug design. *Sci. China Cemet.* **2019**, *62*, 506–514. [CrossRef]
16. Apache RedHawk User Manual. 2019. Available online: <https://pdfcoffee.com/redhawk-pdf-free.html> (accessed on 21 September 2021).