Multi-Feature Data Generation for Design Technology Co-Optimization: A Study on WAT and CP

Shih-Nung Chen[1*] and Shi-Hao Chen[2]

[1*] Department of Information Communication, Asia University, Taichung, Taiwan, R.O.C.

[*]Corresponding email: nung@asia.edu.tw

[2] DigWise Technology Corporation, Ltd., Hsinchu, Taiwan, R.O.C.

hock.chen@digwise-tech.com

Abstract. This study explores the use of Generative Adversarial Networks (GANs) to generate wafer-level Wafer Acceptance Test (WAT) and Chip Probe (CP) test data in semiconductor manufacturing processes, and their application in relevant process and Design-Technology Co-Optimization (DTCO). The generated virtual silicon data includes device performance, physical-electrical characteristics, distribution of wafer process parameters, and implicit information on wafer-level features such as uniformity and defects. This approach enables interdisciplinary teams to overcome data acquisition barriers while ensuring data confidentiality, and it holds significant potential for the development of advanced Electronic Design Automation (EDA) tools in co-optimizing process and chip design flows.

Keywords: Generative Adversarial Network (GAN), Wafer Acceptance Test (WAT), Chip Probe (CP), Design-Technology Co-Optimization (DTCO), virtual silicon data, Electronic Design Automation (EDA).

1 Introduction

The Design-Technology Co-Optimization (DTCO) methodology has been widely discussed and applied in physical design processes to enhance the overall productivity and competitiveness of semiconductor chips. As shown in Fig. 1, it can be analogized to a massive neural network optimization process. Our focus is on optimizing productivity through inference, which includes chip monitoring, WAT-CP-SLT testing, feature correlation analysis, machine learning, and binning strategies with compensation techniques, among others. On the other hand, in the back-propagation optimization phase, we concentrate on optimizing design and process recipes, which encompass chip model calibration using actual measurements, process parameters tuning, timing extraction based on WAT measurements for the device library, customization and optimization of the device library, On-Chip Variation (OCV) regression for local variations, and optimization of design margins and sign-off strategies.

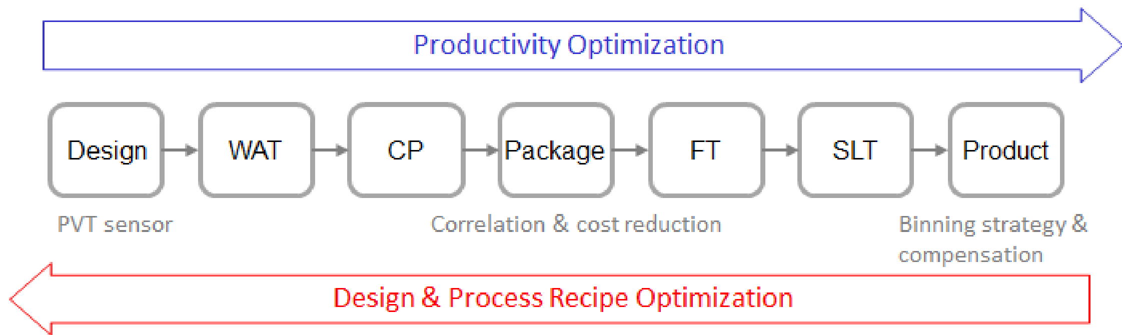

Figure 1. Design-Technology Co-Optimization (DTCO)

However, obtaining and exchanging valuable test data is often challenging and poses barriers to the overall advancement of industry technologies. Therefore, this study proposes an innovative virtual silicon technology that leverages deep learning models to rapidly and accurately generate a large amount of chip data, while reflecting the parameter distribution, defects, and features in wafer manufacturing processes. This study introduces a GAN-based approach that trains and encapsulates multidimensional WAT and CP test data using compact GAN models to generate highly realistic chip data with multidimensional features. This technology plays a crucial role in optimizing chip design and improving the manufacturing process. It brings significant benefits such as enhancing production efficiency, reducing costs, and improving product quality.

2 Background

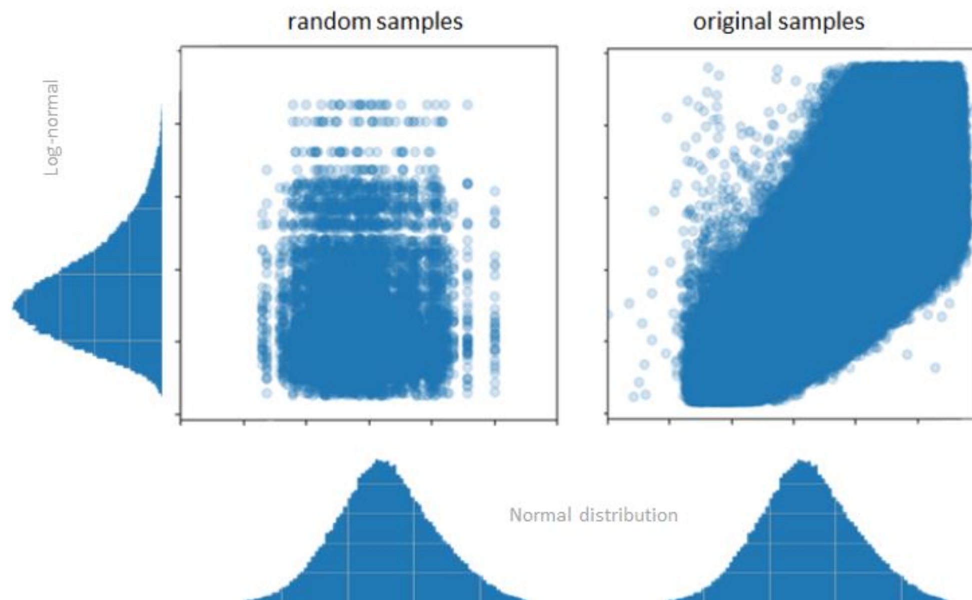

Figure 2. Interdependence and Correlation Among Features

Traditional models often simplify events by assuming Gaussian distributions, disregarding the fact that physical quantities in real chips often exhibit skewed-normal or log-normal distributions. Additionally, these models frequently overlook the interdependencies among vectors in high-dimensional spaces. As shown in Fig. 2, even if each dimension's feature follows the distribution of the parent population when projected individually, the

combined distribution in high-dimensional space may lose the interrelationship between them, somewhat akin to rolling dice.

Due to the multitude of process parameters involved in wafer manufacturing, the relationship between process parameters and wafer or chip-level test data becomes highly intricate, making it challenging for traditional methods to effectively model and analyze. As shown in Fig. 3, even with a comprehensive understanding of the distribution and interrelationship of chip-level features, the lack of wafer-level coordinate information results in the loss of characteristics related to the actual wafer fabrication uniformity. This is a prevalent issue in current simulation analysis modeling. In fact, the distribution of feature vectors in high-dimensional space lacks authenticity, leading to significant discrepancies between production data and simulated data.

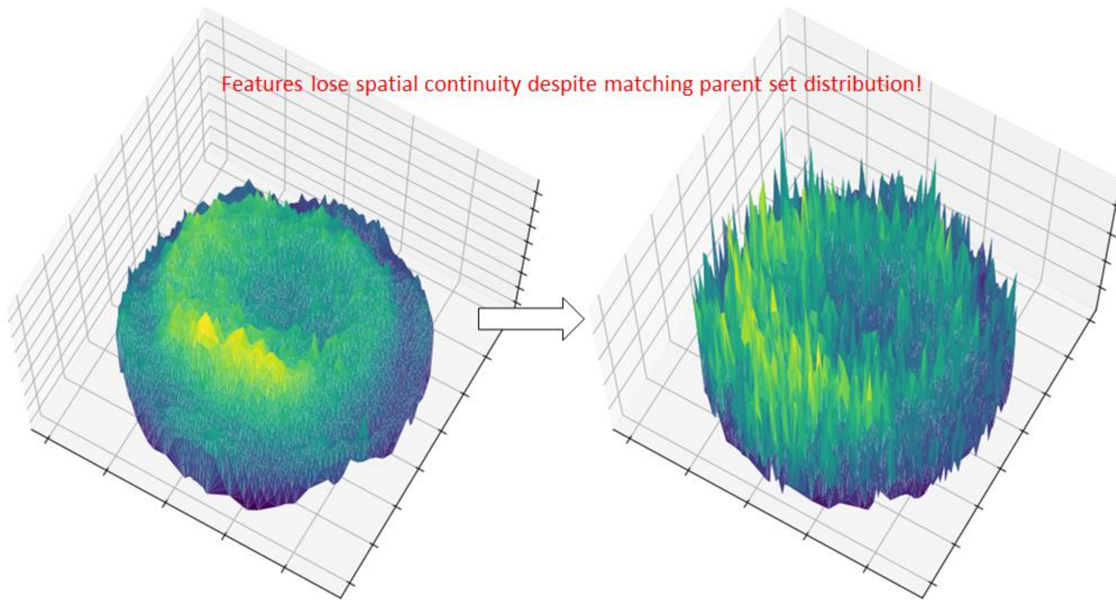

Figure 3. Distribution of Physical Features of Chips at the Wafer-Level Losing Realism

Existing literature has explored the use of deep learning models for simulating and predicting wafer manufacturing processes. Studies [1] and [2] propose a method for wafer defect detection based on a Deep Convolutional Generative Adversarial Network (DCGAN). This approach utilizes a DCGAN to learn the distribution of defect images on wafers and employs the generated model for defect detection and classification. Furthermore, studies such as [3], [4], [5], and [6] demonstrate the potential of Generative Adversarial Networks (GANs) for various other applications in the manufacturing domain.

However, our research differs from existing literature in several aspects. We have successfully achieved the generation of highly realistic virtual silicon data and proposed a platform for chip and wafer-level data analysis and co-optimization based on GAN models. To better capture the variations in wafer-level processes, we have incorporated additional physical features into the construction of the training dataset.

3 System Architecture

This study employs a GAN model to capture the uniformity characteristics of defects and parameters in the wafer

manufacturing process using a large volume of multi-dimensional data. Firstly, we transform the original multi-dimensional data into two-dimensional images, and set multiple feature dimensions (parameter C), as shown in Fig. 4. The size of parameter C is correlated with the network size, and the training time exhibits non-linear growth. Based on computations performed on a personal computer CPU, we select the feature dimensionality C to be between 10 and 18.

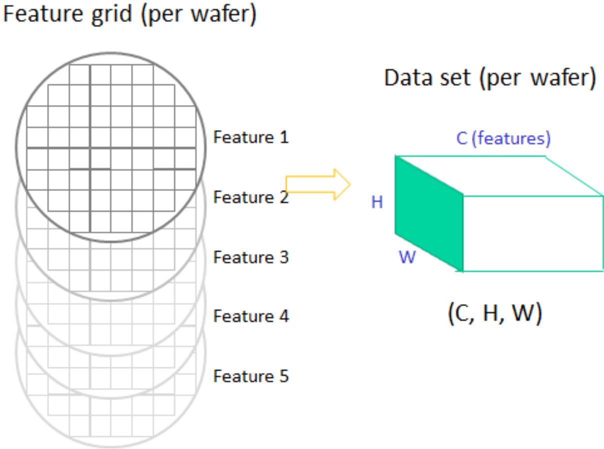

Figure 4. Transformation and Integration of Wafer-Level Multidimensional Training Dataset

In this study, we further augment the training dataset using small angle rotation transformations to simulate the occurrence of rotational defects and process parameters distributions in the wafer manufacturing process, as shown in Fig. 5. This approach enables us to accurately capture the key features in the wafer manufacturing process, thereby improving the training effectiveness of the model.

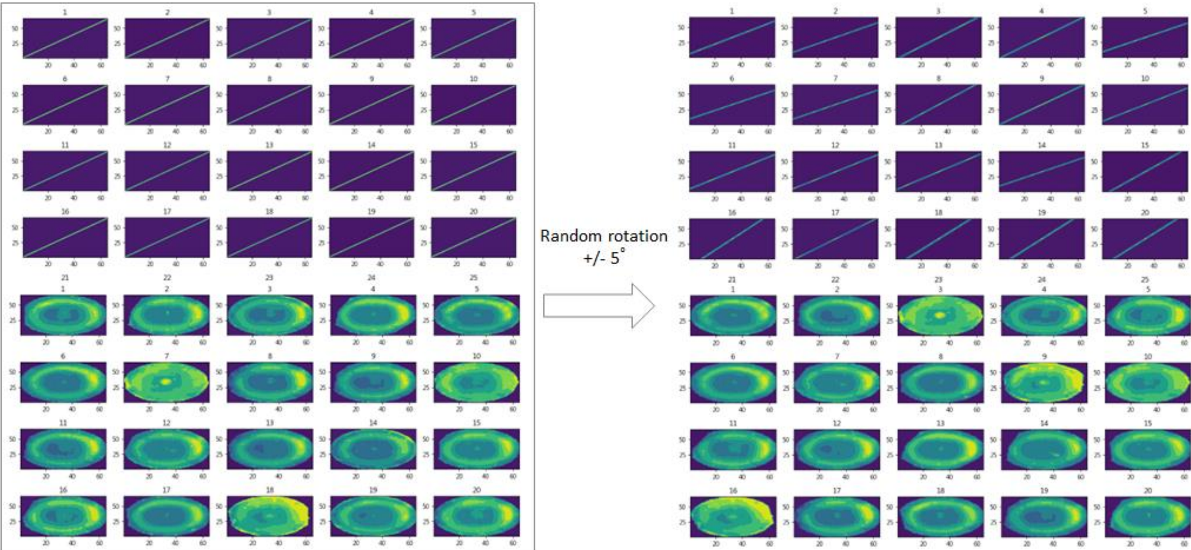

Figure 5. Augmentation of Training Dataset through Small Angle Rotation Transformations

In our study, we utilized a Convolutional Neural Network (CNN) to construct a Generative Adversarial Network (GAN) model, as shown in Fig. 6, for generating chip data with various process features while

incorporating potential defects. The generator component of the model consists of multiple convolutional layers and Tanh activation layers to generate wafer images. Simultaneously, the discriminator component also includes multiple convolutional layers and Sigmoid activation layers to distinguish between real and generated data. These design components work together to achieve the goal of generating high-quality silicon data.

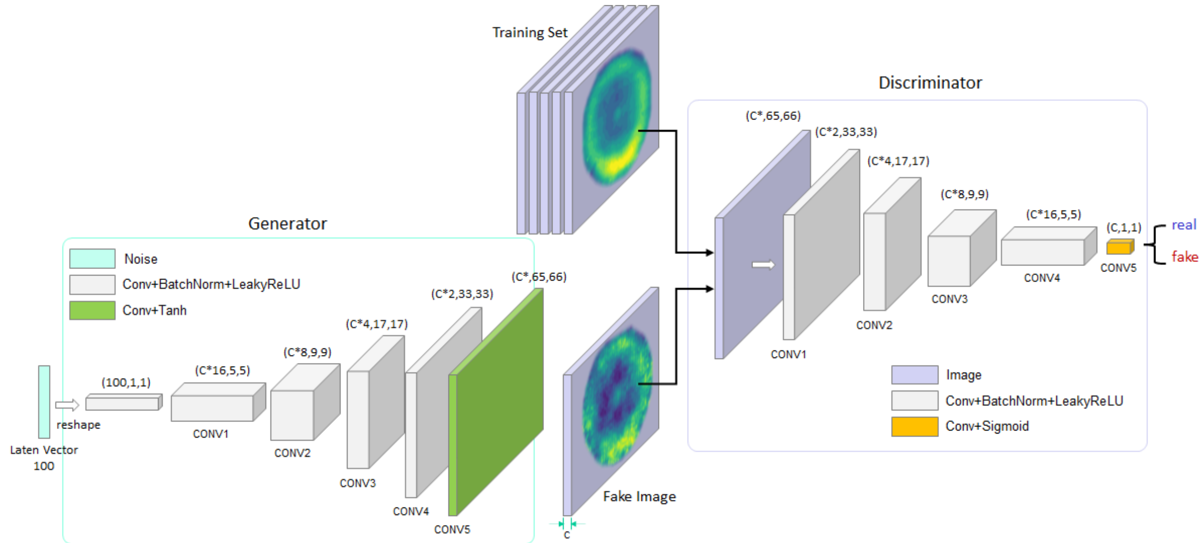

Figure 6. GAN Model

During the model training, we utilized the gradient descent optimization algorithm to minimize the difference between the generated chips and real chips. To enhance the stability of the model, we employed techniques such as batch normalization and the LeakyReLU activation function. Through several hundred iterations, our GAN model was capable of generating highly realistic silicon data, including the chip's position on the wafer, the uniformity of physical features at the wafer level, and the defects present in the chip manufacturing process. These training outcomes provide a reliable data foundation for simulating and analyzing the chip manufacturing process.

4 Experimental Results and Collaborative Optimization Platform

This section will showcase the generated chip data using the GAN model and conduct a detailed analysis, while establishing a Design-Technology Co-Optimization (DTCO) platform. Our dataset consists of approximately 12 million chip data points, with the exclusion of 3σ outliers and missing chip data, deliberately retaining chips with uniformity defects as the training set for the GAN model. In addition to visualizing the data, we also utilize quantitative metrics to evaluate the quality of the generated silicon data. For instance, we use Jensen-Shannon Divergence to compare the similarity of probability distributions between the generated data and real chip data. Additionally, we leverage the Kernel Density Estimation (KDE) metric to quantify the numerical differences between probability distributions of different features. These evaluation methods ensure a reliable and accurate assessment of the generated silicon data quality.

To protect the confidentiality of chip technology and wafer process data, we have uniformly normalized the charts and figures of our research results, limiting the numerical range between 0 and 1.

The experimental results show that the scatter plots between the features of the generated silicon data by the GAN model and the real chip data are highly similar, capturing the process adjustment and variability in the early stages, as shown in Fig. 7. Further analysis using the Jensen-Shannon Divergence index reveals that the probability distributions of the generated silicon data closely align with the characteristics of the real data, ranging from 0.98 to 1.0 across different dimensions, as shown in Fig. 8.

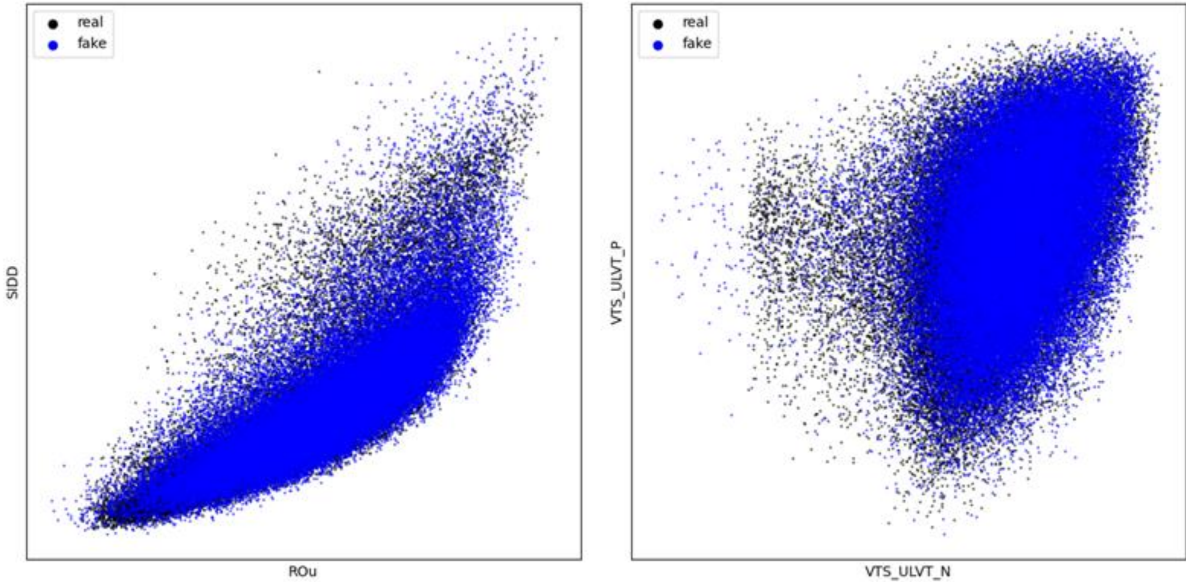

Figure 7. Scatter Plots of Features for Generated Data and Real Data

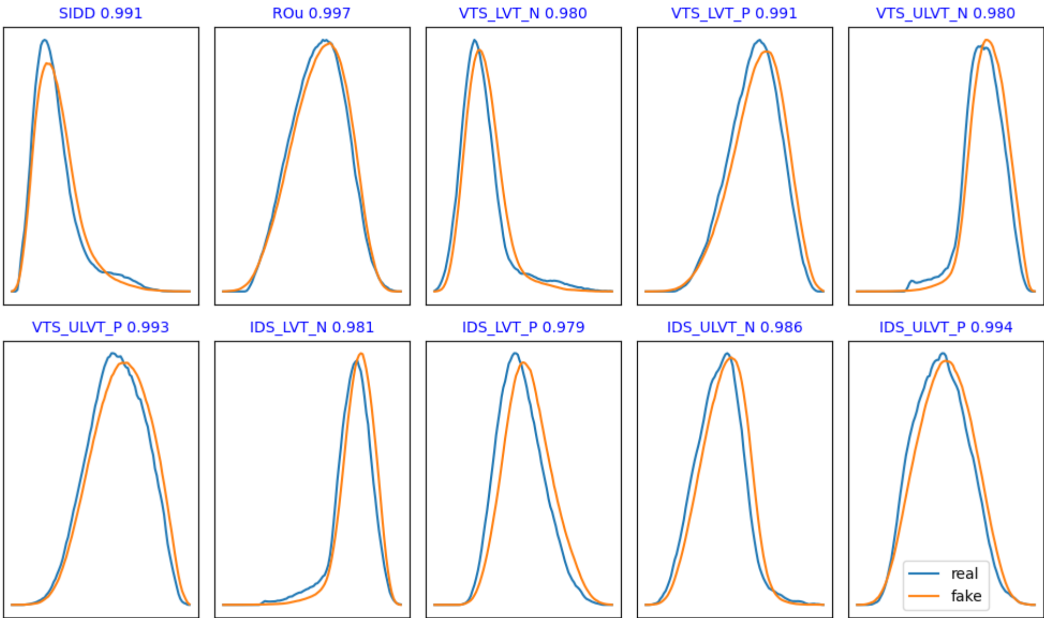

Figure 8. Probability Distribution of Generated Data and Real Data for Each Feature

Furthermore, the combination of generated data in high-dimensional space still preserves the correlations of the original parent population, as shown in Fig. 9.

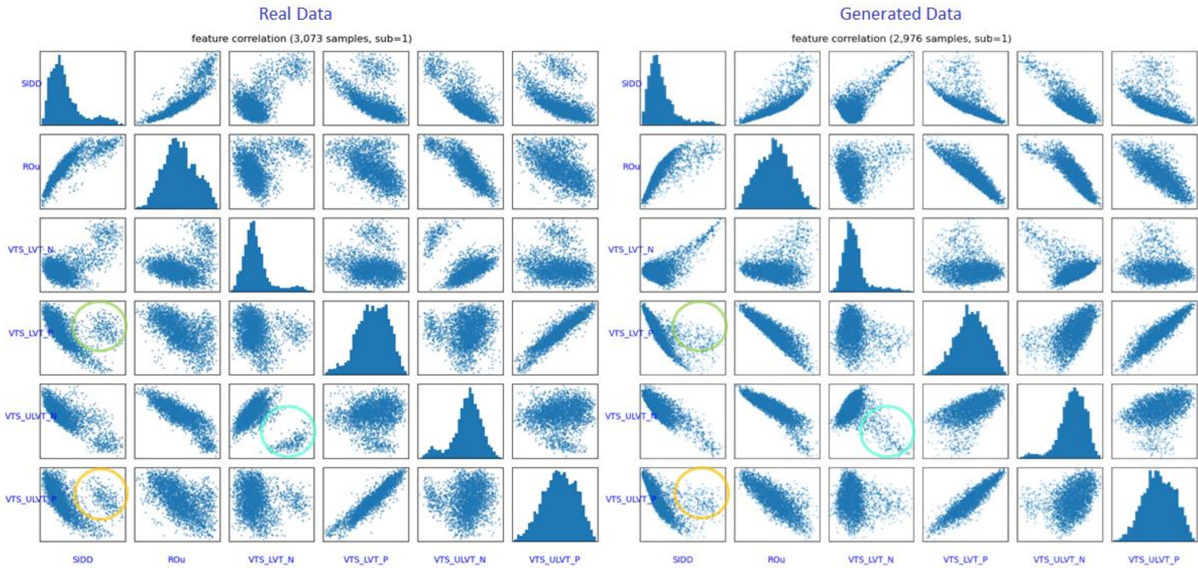

Figure 9. Preservation of Feature Combination in High-Dimensional Space

Figure 10 shows the compromise space between yield and design margin based on the generated large dataset of chip data, providing specific guidance for future design recipes and capacity optimization. The model demonstrates good stability and generalization performance across different training and testing sets.

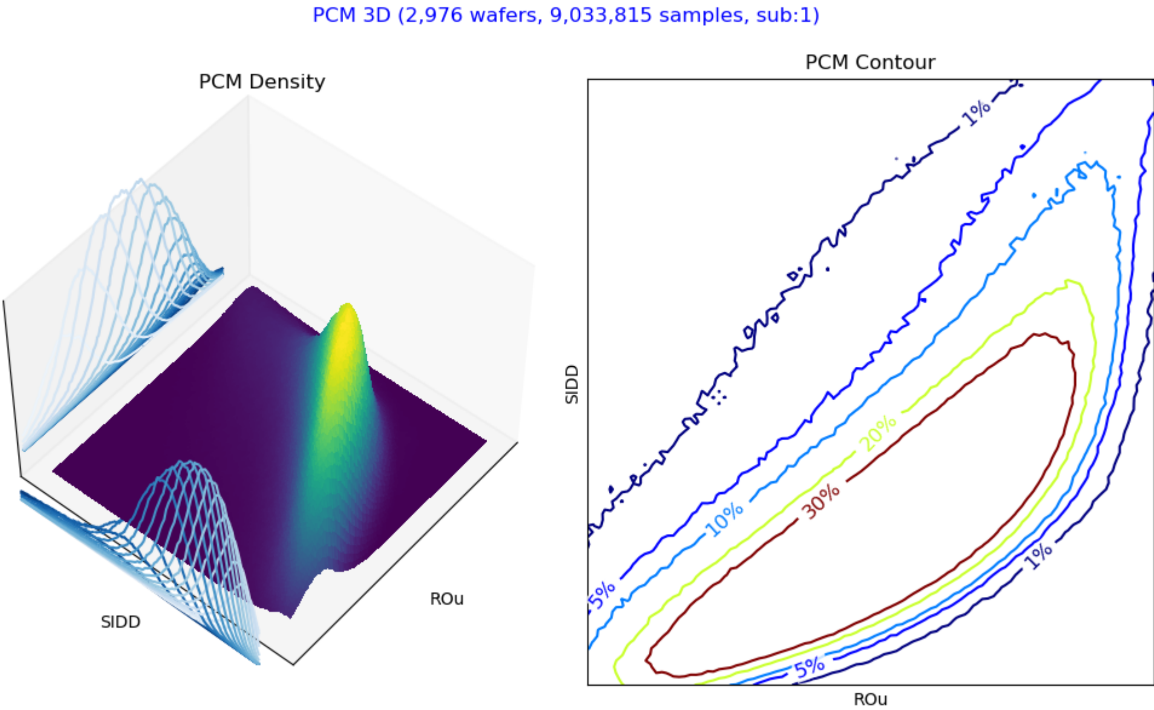

Figure 10. Probability Density Distribution of Multidimensional Features

5 Conclusions

This study presents the application of Generative Adversarial Networks (GANs) in chip and wafer test data modeling and silicon virtualization. The study explores the utilization of virtual silicon data in Design-Technology Co-Optimization (DTCO) and showcases several related design and process co-optimization schemes. The research aims to assist process and chip design engineers in generating more realistic design examples, facilitating trade-offs between different process recipes and binning strategies for overall capacity optimization. Additionally, the study contributes to the optimization of process parameters and design margins, enabling better energy efficiency designs. However, GANs also face challenges, including selecting appropriate generator and discriminator architectures and handling higher-dimensional and complex data. Furthermore, training the generator and discriminator models requires significant time and computational resources.

To enhance the transformation of data into trainable models, we employ a method that converts multidimensional data into two-dimensional images with multiple feature channels. This approach enables us to simulate uniformity defects and variations in process parameters that may occur in wafer manufacturing during the training process. The research findings substantiate the effectiveness of this proposed method in supporting chip design, product optimization, and process improvement. It leads to enhanced production efficiency, cost reduction, and improved product quality, thereby offering valuable contributions to the industry.

In summary, GANs hold enormous potential for the development of advanced EDA tools. By harnessing GAN-generated models to capture the intricate mapping between process and design, we can achieve enhanced efficiency in process and design optimizations. Nevertheless, further research and development efforts are necessary to address current challenges and limitations. We eagerly anticipate increased attention from researchers in this domain, as they continue to explore and propose innovative solutions for the future.

References

1. Wang, J., et al. AdaBalGAN: An improved generative adversarial network with imbalanced learning for wafer defective pattern recognition. IEEE Transactions on Semiconductor Manufacturing, Vol. 32, Issue 3, pp. 310-319 (2019)
2. Hu, G., et al. Unsupervised fabric defect detection based on a deep convolutional generative adversarial network. Textile Research Journal, Vol. 90, Issue 3-4, pp. 247-270 (2020)
3. Kusiak, A., Convolutional and generative adversarial neural networks in manufacturing. International Journal of Production Research, Vol. 58, No. 5, pp. 1594-1604 (2020)
4. Singh, R., et al. Generative adversarial networks for synthetic defect generation in assembly and test manufacturing. In: 2020 31st Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC). IEEE, pp.1-5 (2020)
5. Hertlein, N., et al. Generative adversarial network for early-stage design flexibility in topology optimization for additive manufacturing. Journal of Manufacturing Systems, Vol. 59, pp. 675-685 (2021)
6. Qin, J., et al. A novel temporal generative adversarial network for electrocardiography anomaly detection, Artificial Intelligence in Medicine, Vol. 136, 102489 (2023)