

# Enhancing an Autoregressive Generative Wafer Polishing U-net Model

Aixtal Corporation<sup>1</sup>, Mipox Corporation<sup>2</sup>, Kevin Richard G. Operiano<sup>1</sup>, Roberto Iaconi<sup>1</sup>,  
Riku Tanaka<sup>1</sup>, Fumiya Kawate<sup>1</sup>, Sepasy Saeed<sup>2</sup>, Yoshifumi Watanabe<sup>2</sup>

E-mail: k.operiano@aixtal.com

## Background

Deep learning models are susceptible to prediction errors when used in simulating manufacturing processes. One of the errors, as we have observed during the prediction of the polishing amount profiles of the SiC wafer and its pad, which requires the use of output profile as a model input recursively, leads to exponential compounding errors caused by *exposure bias*. To mitigate this problem, we propose methods (i.e., dataset preprocessing and post-processing, and autoregressive training) to promote model robustness and output stability.

## Methods

We propose three methods to improve an autoregressive U-net model training and evaluation for SiC wafer polishing. The model inputs are the front and back surface wafer profiles, the upper and lower polishing pad profiles, as well as the simulation parameters (e.g., rotation speed, pressure, etc.). The output is the polishing amount profile (i.e., the profile of the material removed from all inputs).

- **Preprocessing** - Gaussian noise, quantization noise, and impulse noise are randomly introduced as data augmentations.
- **Post-processing** - Before using the model output as an input, the negative values are zeroed (i.e., the negative polish amount is removed).
- **Autoregressive Training** - The model is trained on the loss of its generated output as input over multiple steps, in addition to ground truth.

## Results, Discussion, and Conclusion

Four model configurations are experimented with, namely: vanilla U-net, with preprocessing (PR), with pre- and post-processing (PP), and pre- and post-processing combined with Autoregressive Training (AT). Fig. 1 shows that each proposed method gradually stabilizes the profile prediction to a certain degree. Combining them yields the most robust prediction even after eight polish passes. Preprocessing handles the prediction fluctuations, while post-processing and autoregressive training provide output stability and model resiliency. The proposed methods effectively handle exponential compounding errors caused by recursive model evaluation.

## Acknowledgement

This paper is based on results obtained from a project, JPNP21029, subsidized by the New Energy and Industrial Technology Development Organization (NEDO).

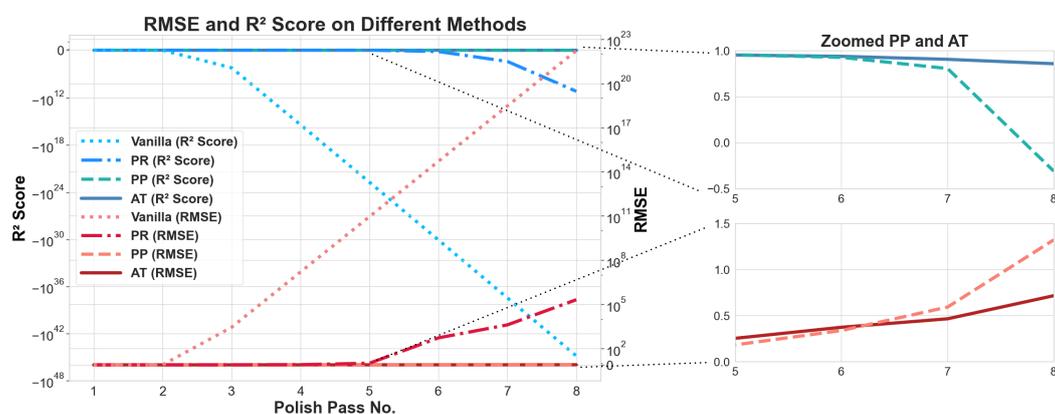


Figure 1: RMSE and R<sup>2</sup> score metrics of various proposed techniques on consecutive polish passes