PriMera Scientific Engineering
Volume 9 Issue 1 July 2026
DOI: 10.56831/PSEN-09-279
ISSN: 2834-2550

PriMera Scientific Publications

Integrated Hierarchical Reinforcement Learning for Enhanced CMP Manufacturing Efficiency

Type: Research Article
Received: June 02, 2026
Published: June 30, 2026

Citation:
Andy Yendi Tsen. "Integrated Hierarchical Reinforcement Learning for Enhanced CMP Manufacturing Efficiency". PriMera Scientific Engineering 9.1 (2026): 11-21.

Andy Yendi Tsen*

Department of Industrial Management, National Taiwan University of Science and Technology, Taiwan

***Corresponding Author:** Andy Yendi Tsen, National Taiwan University of Science and Technology, Taipei City, Taiwan.

Abstract

This study demonstrates that applying artificial intelligence (AI) to chemical-mechanical planarization (CMP) process control significantly reduces rework rates and enhances manufacturing efficiency. We propose a hierarchical AI CMP controller featuring an auto-tuning capability. This integrated hierarchical reinforcement learning (iHRL) framework is designed to mitigate CMP variations arising from removal rate decay and pattern density fluctuations. The iHRL agent utilizes a structured wafer-to-wafer and lot-to-lot control architecture. Results demonstrate that the self-learning auto-tuning mechanism effectively minimizes the CMP rework rate. Furthermore, comprehensive simulations indicate that the proposed controller enhances average manufacturing efficiency by more than double digits compared to the conventional run-to-run controller.

Keywords: semiconductor manufacturing; chemical mechanical planarization; artificial intelligence

Introduction

Manufacturing efficiency in semiconductor fabrication remains a significant challenge due to the continuous emergence of advanced semiconductor technologies. The complexity of semiconductor manufacturing arises not only from sophisticated fabrication processes but also from the diverse physical properties of silicon chips, which vary across the versatile integrated circuit (IC) products of different designers. Chemical-mechanical planarization (CMP) is a critical fabrication process significantly impacted by both equipment decay during wafer processing and pattern density variations resulting from different IC designs. In foundry-based semiconductor manufacturing, run-to-run (R2R) control has been employed for decades to compensate for CMP equipment decay. Concurrently, operators group wafers with identical IC designs into specific lots and apply recipes based on heuristic expertise. This approach aims to control both wafer-to-wafer and lot-to-lot variations simultaneously to ensure high yield and manufacturing efficiency. However, the inherent delay in post-CMP measurements remains a major constraint, limiting the potential for further improvements in manufacturing efficiency.

Background and related works

This study applies integrated hierarchical reinforcement learning (iHRL) to enable auto-tuning capabilities for CMP wafer-to-wafer and lot-to-lot control, even in the absence of real-time metrology feedback. This application of artificial intelligence reveals the potential for achieving optimal manufacturing efficiency.

Run-to-run control

In semiconductor manufacturing, the gradual degradation of CMP polishing materials and equipment is strongly correlated with both wafer fabrication quality and manufacturing efficiency. It is essential to compensate for this variation—specifically, the removal rate decay— by adjusting polishing recipes across different wafers. The model predictive control (MPC) [1] strategy is frequently employed to optimize automatic CMP recipe tuning, commonly referred to as run-to-run (RtR) control. Consequently, RtR control utilizing specific stochastic filters, such as the exponentially weighted moving average (EWMA), has been designed and implemented in real-world scenarios for decades [2-5].

The application of artificial intelligence (AI) to enhance semiconductor process control is advancing rapidly. AI, particularly through neural networks, has become a prominent pathway for manufacturing optimization and is widely applied to discrete process control [6-8]. The optimal design of artificial neural network controllers has been extensively investigated [9]. In the semiconductor industry, various neural network models and deep reinforcement learning (DRL) approaches are being integrated with semiconductor run-to-run (RtR) control. Additionally, reinforcement learning applications are being extensively implemented across semiconductor manufacturing [10, 11]. The complexity of manufacturing challenges and the requirement for coordinated multi-agent actions have driven advanced research into hierarchical reinforcement learning (HRL) [12-14].

Wafer pattern density

In a semiconductor foundry (FAB), the variety of processed wafer types—each characterized by a specific IC design and pattern density—can number in the hundreds. The volume of material removed over a fixed polishing duration at any given location, which correlates with the proportion of raised topography on the wafer die, is defined as the effective pattern density. The feature size and pattern density of the designed die influence localized pressure distribution, thereby affecting the removal rate. Physically, die areas with lower pattern densities polish at a faster rate than those with higher pattern densities [15]. In operational practice, wafers with similar pattern densities are grouped into the same lot. CMP control must simultaneously address wafer-to-wafer removal rate decay and lot-to-lot disturbances stemming from pattern density variations. Therefore, developing a reliable, high-performance model for effective CMP control is crucial.

Feature size and pattern density fundamentally affect the mechanical and chemical polishing mechanisms in CMP. Some studies have proposed building CMP models based on the dependence of local polishing effects [16], stating that the material removal rate is determined by the differential down-pressure between different topographic regions. One well-known CMP modeling approach is based on systematic experimental design [17]. Research analyzing the fundamental effects of pattern density on the CMP process has shown that existing CMP models can be applied not only to local regions but also to the global area of dies on wafers. Furthermore, efficient chip-level modeling of CMP wafer thickness across arbitrary product die patterns has also been explored [18, 19].

CMP manufacturing performance

The chemical and physical polishing stresses applied to the wafer surface during the CMP process pose a significant risk of wafer breakage or scrapping. Consequently, CMP typically involves a high-speed polishing stage followed by a low-speed polishing stage on the same wafer. Following the first stage, engineers reduce the polishing speed and set the final polishing time for the second stage based on metrology measurements. This repeated polishing process on the same wafer is referred to as “CMP rework”. If the initial polish successfully meets the target thickness, the second stage can be bypassed; however, if not, the rework operation becomes

time-consuming and labor-intensive. Minimizing rework is a primary challenge in enhancing CMP manufacturing efficiency.

In CMP fabrication, it is particularly challenging to determine in-situ when the target material removal has been achieved or when the optimal polishing degree has been reached. This difficulty necessitates additional measurements to verify process completion and minimize rework. Ultimately, process performance directly dictates the frequency of reworked wafers. The histogram of wafer-mean film thickness under CMP control serves as an indicator for achieving higher wafer throughput while maintaining low scrapping risks. As illustrated in figure 1, increased variation in the control chart directly correlates with a higher CMP rework rate.

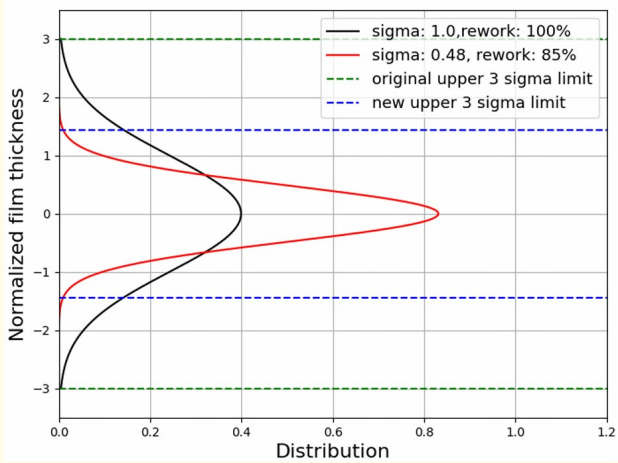

Figure 1: The histogram of wafer-mean film thickness with different CMP rework rates.

Materials and Methods

Wafer-to-wafer and lot-to-lot CMP control

In practical semiconductor manufacturing, the execution of CMP (chemical-mechanical planarization) control must account for both wafer-to-wafer removal rate decay and the compensation for lot-to-lot variations arising from diverse intrinsic pattern density changes. The proposed control scenario is illustrated in the control block diagram presented in figure 2.

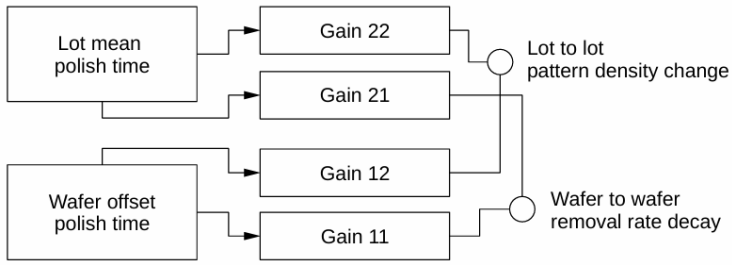

Figure 2: Wafer-to-wafer and lot-to-lot control block diagram.

It is proposed to assign the polish time as the manipulated variable of the CMP controller. The CMP polish time is quantified as the lot-mean polish time and wafer-offset polish time, as shown in the following equations.

$$t_{lot-mean} = \frac{\sum_{i}^{N_{wafer}} t_{i,polish}}{N_{wafer}} \quad (1)$$

$$t_{i,offset} = t_{i,polish} + t_{lot-mean} \quad (2)$$

Where $t_{lot-mean}$ is the average lot polish time, N_{wafer} is the number of wafers per lot, $t_{t,polish}$ is the polish time on specific wafer; $t_{i,offset}$ is the wafer-offset polish time on specific wafer, $t_{ipolish}$ is the wafer polish time on specific wafer.

As shown in Figure 2, the interactions between the manipulated variables and the controlled variables are denoted by G_{12} and G_{21}, respectively. To implement an effective hierarchical CMP control strategy, it is essential to decouple these multi-variable interactions. Specifically, the interaction gains, G_{12} and G_{21}, must be minimized to ensure the successful implementation of this hierarchical multi-variable controller.

In practical semiconductor manufacturing, FAB engineers group wafers with identical pattern densities into the same lots. This practice ensures that there is no wafer-to-wafer pattern density variation within a single lot, thereby driving the value of G_{21} toward zero. Conversely, the impact of recipe changes are the same on all wafrers—driven by lot-to-lot pattern density variations—should be eliminated during lot processing. We assume that the interaction between the wafer-offset polishing time and the lot-to-lot polishing time is negligible, such that the pattern density-induced interaction (G_{12}) approaches zero, as implied by the constraint in (3).

$$\sum_{i}^{N_{wafer}} t_{i,offset} \cong 0 \quad (3)$$

As illustrated in figure 3, the decoupled control loop is implemented within a hierarchical architecture designed for both wafer-to-wafer and lot-to-lot process control. On the right side of the figure, during the wafer treatment phase of the CMP process, a portion of wafers may require rework if their film thickness exceeds specified limits. Such rework operations incur additional costs and processing time, which negatively impact manufacturing throughput and quality. On the left side of the figure, the hierarchical controller learns the dynamics of both pattern density variations and removal rate decay. Specifically, the controller predicts optimal control actions to mitigate lot-to-lot and wafer-to-wafer variations in the CMP process. This architecture enables minimal process variation and reduced rework rates, thereby achieving optimal manufacturing efficiency. Simultaneously, a reinforcement learning algorithm, leveraging artificial intelligence, operates alongside the hierarchical predictive controller to continuously explore and learn the underlying patterns of pattern density changes and removal rate decay.

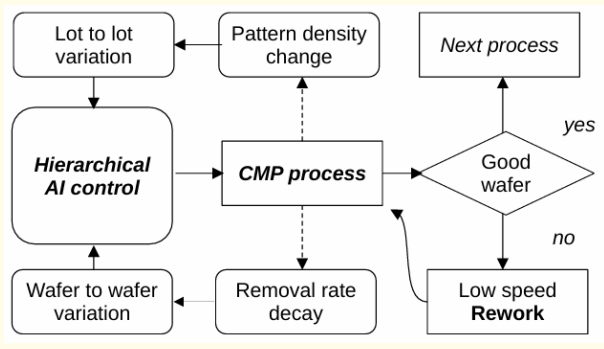

Figure 3: Hierarchical lot-to-lot and wafer-to-wafer control.

Integrated hierarchical reinforcement learning control

As illustrated in Figure 4, the proposed iHRL CMP controller is implemented on a simulation platform. This study adopts the Deep Deterministic Policy Gradient (DDPG) reinforcement learning algorithm. The architecture employs high-level lot-to-lot and low-level wafer-to-wafer strategies to simultaneously compensate for pattern density-induced variations and removal rate decay, as depicted in Figure 4. The implemented DDPG mechanism comprises the following components:

Action Space

The action space for the low-level strategy consists of the CMP wafer-offset polishing time ($t_{i,offset}$) defined in (2), which is utilized to compensate for the removal rate decay on each individual wafer. The action space for the high-level strategy is the lot-mean polishing time ($t_{lot-mean}$) defined in (1), which compensates for the pattern density variations across different lots. The action space of this hierarchical DDPG framework is subject to the physical constraints specified in (3).

Reward Functions

The reward functions are designed based on the negative absolute error between the post-CMP film thickness and the target thickness. Specifically, the high-level strategy reward is defined as the negative absolute difference between the measured lot-mean film thickness and the target lot thickness. The low-level strategy reward is defined as the negative absolute difference between the measured wafer film thickness and the predicted lot-mean thickness. Since wafers are measured in a sequential manner, the low-level reward is evaluated using the measured wafer film thickness relative to the predicted lot-mean thickness provided by the high-level strategy.

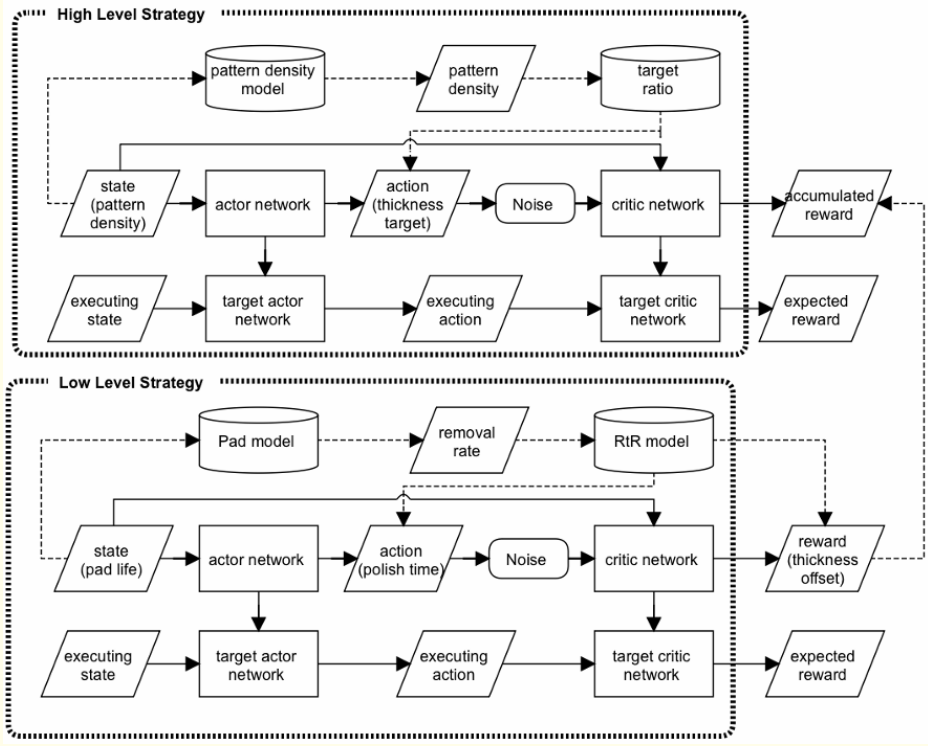

Figure 4: Integrated hierarchical reinforcement learning CMP control.

Required auto-tuning features of CMP controller

In theory, if the post-CMP wafer film thickness were perfectly predictable, the CMP process could be controlled with absolute precision. However, in practical applications, measurement delays and prediction errors are inevitable. To illustrate this, consider a scenario involving 100 wafers divided into four lots, subject to a 5% pattern density variation between lots and a six-wafer measurement delay per lot. The film thickness dynamics for these 100 wafers are simulated and presented in figure 5. The black line represents the uncontrolled dynamics. When engineers manually adjust the CMP recipes based on varying pattern densities, the resulting dynamics are depicted by the green dashed line. In this simulation, the pattern density values for the four lots are assigned as 0.72, 0.76, 0.84 and 0.88. When the conventional RtR-EWMA control is applied, the dynamics converge to the blue line, which serves as the baseline for modern CMP performance.

In large-scale semiconductor manufacturing, implementing real-time film thickness metrology without measurement delay is infeasible. Consequently, the period during which the controller operates without measurement feedback is termed the "***blind horizon***". as indicated by the dashed intervals in figure 5. While the original RtR-EWMA control aims to mitigate the impact of this "***blind horizon***" using statistical techniques, the proposed iHRL CMP control seeks to eliminate variations arising from pattern density changes and removal rate decay during this period by leveraging learned intelligence from artificial intelligence.

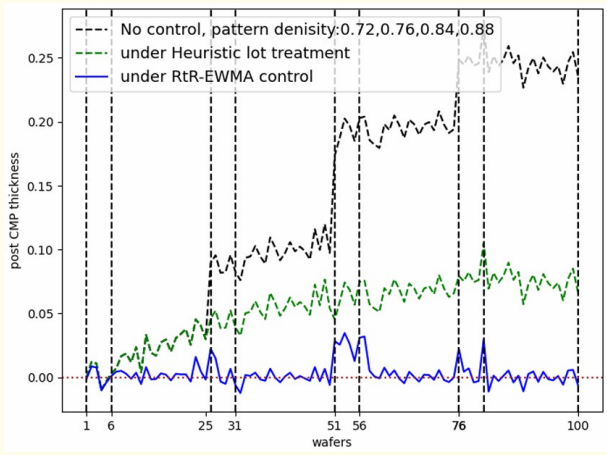

Figure 5: CMP wafer thickness dynamics with different treatment scenario.

Results and Discussion

Wafer-to-wafer control: consistent pattern density change

The proposed hierarchical reinforcement learning (iHRL) CMP controller is implemented as illustrated in figure 4. This architecture incorporates a simulated CMP model that captures the dynamic interactions between film thickness and removal rate decay with defined patter density during lots. The experimental data used for the model-fitting process are derived from the work of Stine [17]. The performance of the iHRL mechanism is evaluated through these simulation results. To address the "***blind horizon***" in CMP control, the system operates in two modes: when real-time measurement feedback is available, the process controller follows the conventional algorithm, namely the RtR-EWMA control; however, during measurement delays (i.e., when feedback is unavailable), the iHRL controller compensates for both wafer-to-wafer and lot-to-lot variations by leveraging knowledge from the CMP model and the reward signals derived from film thickness measurements.

To validate the controller, a simulation involving 100 wafers distributed across four lots was conducted, subject to a 5% lot-to-lot pattern density variation and a six-wafer measurement delay, as shown in figure 6. The green dashed line represents the outcome of the heuristic treatment employed by engineers. In this scenario, the standard deviation of the film thickness for the 100 wafers is 0.0241. The blue line represents the performance of the original RtR-EWMA control, achieving a reduced standard deviation of 0.0094. The red line depicts the performance of the proposed iHRL control, which further reduces the standard deviation to 0.0075. Compared to the heuristic approach used by engineers, both the RtR-EWMA control and the proposed iHRL control reduce process variation. The impacts of process variation reduction are analyzed in the next paragraph.

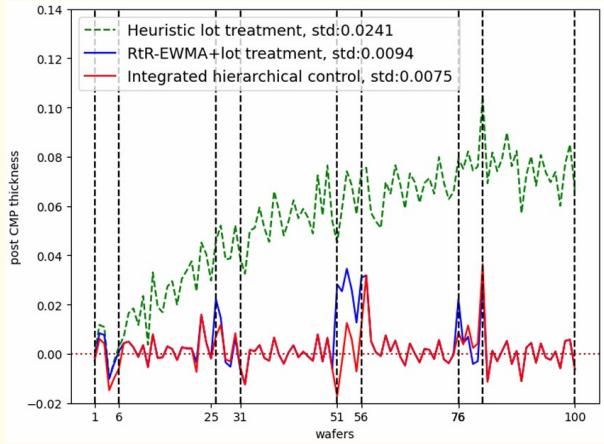

Figure 6: CMP wafer-to-wafer control results.

Performance index: rework rate and manufacturing efficiency

The CMP rework operation aims to manage the risk of wafer breakage or scrapping, but it impacts manufacturing efficiency. The rework rate (r_{rework}) can be defined as (4), where N_{wafer} is the total wafer count and N_dura is the count of wafers which are dual processed. The CMP manufacturing efficiency is defined as wafer outputs per CMP process as shown in (5). In the condition of 100% wafer reworked, the CMP manufacturing efficiency (Eff_{CMP}) is 0.5.

$$r_{rework} = \frac{N_{dual}}{N_{wafer}} \tag{4}$$

$$Eff_{CMP} = \frac{1}{1 + r_{rework}} \tag{5}$$

The rework rate is directly correlated with the distribution of polished wafer thickness; specifically, higher process performance leads to a lower rework rate. In a baseline scenario without any CMP controller, the rework rate is effectively 100%, where the accumulated area of the wafer count histogram from -3 σ to +3 σ represents the entire population, as shown by the black line in figure 7. Both the RtR-EWMA and iDL-HRL CMP controllers are designed to reduce process variation, thereby decreasing the number of wafers requiring rework. The reduction in reworked wafers is represented by the gray area in figure 7. Under the proposed control strategies, the rework rate for the RtR-EWMA controller is reduced to 75.84% (Eff_{CMP}: 56.9%), whereas the iHRL controller further reduces it to 64.98% (Eff_{CMP}: 60.6%).

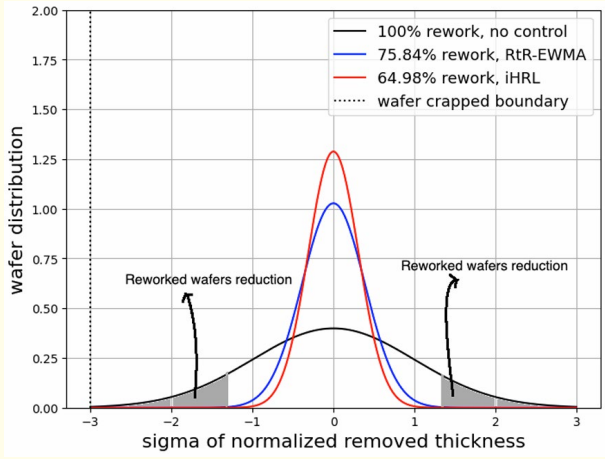

Figure 7: The reduced rework wafers by controllers.

Conversely, as process performance is enhanced, an opportunity arises to tighten the control limits for wafer thickness specifications, a strategy referred to as "target shifting". The target of the CMP controller, which corresponds to the mean value of the thickness distribution, can be shifted, as illustrated in Figure 8. This shift leads to greater manufacturing efficiency, as it reduces the amount of silicon material that must be removed. Under this configuration, the rework rate for the RtR-EWMA controller is 24.97% (Eff_{CMP}: 80.0%), whereas the iHRL controller achieves a significantly lower rate of 12.54% (Eff_{CMP}: 88.9%). Consequently, the proposed iHRL control reduces the rework rate by 12.46% compared to the original RtR-EWMA CMP controller.

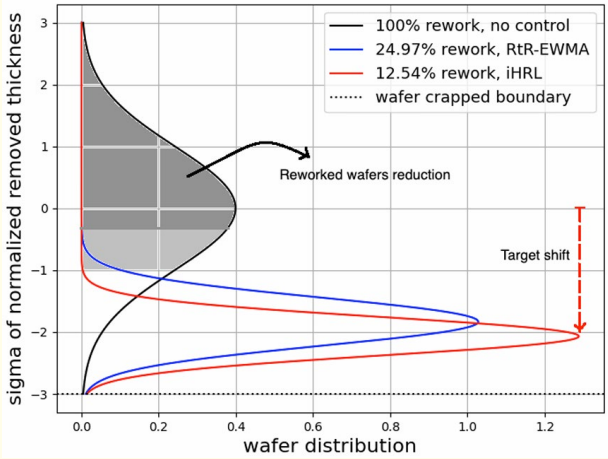

Figure 8: Reduced rework wafers by CMP target shift and controllers.

Total control performance: exhausting pattern density change cases

The performance of various scenarios, characterized by different lot sequences and pattern densities, was investigated to evaluate the impact of lot-to-lot pattern density variations. The simulation involved 100 wafers undergoing the CMP process, with pattern densities varying by 5% per lot. The simulated pattern densities were 0.72, 0.76, 0.84, and 0.88, representing 'very low (VL)', 'low (L)', 'high (H)', and 'very high (VH)' levels, respectively. As previously demonstrated, the process performance and rework rate for the ascending sequence (i.e., 'VL-L-H-VH') were presented in the preceding sections. To provide a comprehensive analysis, the control performance for all 24 possible combinations of these 'VL', 'L', 'H', and 'VH' lot conditions was simulated, as illustrated in figure 9 and table 1.

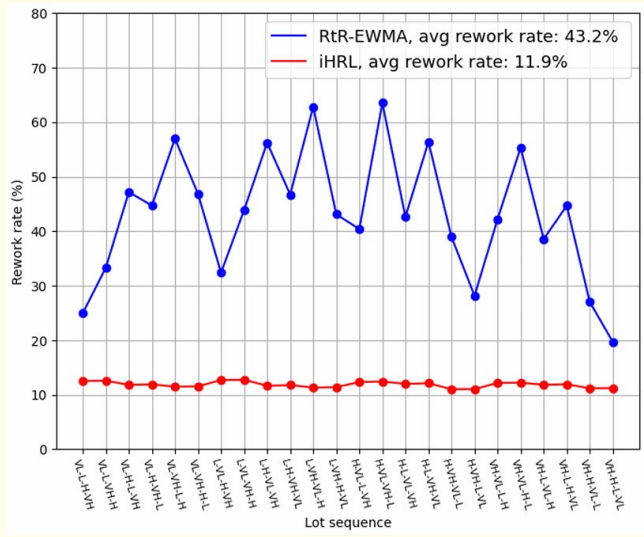

Figure 9: The total performance comparison between iHRL and RtR-EWMA.

	Rework rate				***Efficiency***			
	average	***max***	***min***	***range***	***average***	***max***	***min***	***range***
RtE EWMA	43.2%	63.5%	19.7%	43.8%	70.3%	83.6%	61.2%	22.4%
iHRL	11.9%	12.7%	11.0%	1.7%	89.4%	90.0%	88.7%	1.4%
Enhanced	***31.3%***	50.8%	8.7%	***25x***	***19.1%***	6.4%	27.5%	***16x***

Table 1: Comparisons of rework rate and manufacturing efficiency.

The average performance of the proposed iHRL controller significantly outperforms the conventional RtR-EWMA approach. Specifically, the rework rate is reduced by 31.3%, while manufacturing efficiency is enhanced by 19.1%. As detailed in table 1, the most striking difference between the original RtR-EWMA and the iHRL controller lies in the disparity between their best and worst-case performance ranges. The rework rate for RtR-EWMA fluctuates widely, ranging from 19.7% to 63.5% (a spread of 43.8%). In stark contrast, the iHRL controller maintains a much tighter range, with the rework rate varying only from 11.0% to 12.7% (a spread of less than 2%).

This stability clearly demonstrates that the iHRL framework is far more robust and effective in managing process variations induced by pattern density changes—a critical challenge in foundry fabrication plants. These results provide strong evidence that iHRL is the key to control the lot-to-lot variations in CMP caused by pattern density shifts, whereas the original RtR-EWMA is notably weak in this regard. This disparity suggests that the static filtering mechanism of the RtR-EWMA controller lacks the necessary adaptability under varying pattern density conditions, as such-density fluctuations represent a time-independent variable that cannot be effectively man-

aged by a standard exponentially weighted moving average (EWMA) filter. Conversely, the AI adaptive feature of the iHRL controller achieves superior control performance.

Conclusion

The physical limitation imposed by measurement delay and diversity of wafer pattern density in semiconductor metrology significantly impacts both process performance and manufacturing efficiency. For decades, static control approaches have been employed to mitigate this limitation; however, they struggle to manage the non-linear variations arising from environmental factors and heuristic-based-treatments. With the advancement of artificial intelligence, adaptive control approaches possessing learning capabilities offer a promising pathway toward superior process performance and higher manufacturing efficiency. In this study, the average chemical-mechanical planarization (CMP) manufacturing efficiency is remarkably enhanced by over 19%. Notably, the performance disparity between the best and worst-case scenarios is reduced to less than 2%, representing a 25-fold improvement in rework rate stability and a 16-fold improvement in efficiency stability compared to the original static process controller.

In an advanced semi-conductor manufacturing fabrication plant (FAB), there are more than 10 stages of the CMP process, the overall FAB manufacturing efficiency enhancement is highly anticipated with the implementation of this AI auto-tuning CMP controller.

References

1. A-C Lee, Y-R Pan and M-T Hsieh. "Output disturbance observer structure applied to run-to-run control for semiconductor manufacturing". IEEE Transactions on Semiconductor Manufacturing 24.1 (2011): 27-43.
2. WJ Campbell., et al. "A comparison of run-to-run control algorithms". in Proceedings of the 2002 American Control Conference (IEEE Cat. No. CH37301) IEEE 3 (2002): 2150-2155.
3. AC Lee, TW Kuo and ZL Lee. "Modified double EWMA approach for mixed product run-to-run CMP process control". Advanced Materials Research 314-316 (2011): 2504-2511.
4. E Sachs, A Hu and A Ingolfsson. "Run by run process control: Combining SPC and feedback control". IEEE Transactions on Semiconductor Manufacturing 8.1 (1995): 26-43.
5. Y Zheng., et al. "Model quality evaluation in semiconductor manufacturing process with EWMA run-to-run control". IEEE Transactions on Semiconductor Manufacturing 30.1 (2017): 8-16.
6. TP Lillicrap., et al. "Continuous control with deep reinforcement learning". arXiv preprint arXiv:1509.02971, (2015).
7. MM Thomas, B Joseph and JL Kardos. "Batch chemical process quality control applied to curing of composite materials". AIChE journal 43.10 (1997): 2535-2545.
8. AYD Tsen., et al. "Predictive control of quality in batch polymerization using hybrid ANN models". AIChE Journal 42.2 (1996): 455-465.
9. TH Smith and DS Boning. "A self-tuning EWMA controller utilizing artificial neural network function approximation techniques". in IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part C 20.2 (1997): 121-132.
10. Y Li, J Du and W Jiang. "Reinforcement learning for process control with application in semiconductor manufacturing". IISE Transactions 56.6 (2024): 585-599.
11. A Yen-Di Tsen and T-L Chen. "Reinforcement learning chemical-mechanical polishing run-to-run controller". in 2023 IEEE 5th Eurasia Conference on IOT, Communication and Engineering (ECICE). IEEE (2023): 732-735.
12. AY Tsen, T-L Chen and K-S Hsu. "Deep Hierarchical Reinforcement Learning CMP Run to Run Control". in International Conference on Flexible Automation and Intelligent Manufacturing. Springer (2024): 58-65.
13. AG Barto and S Mahadevan. "Recent advances in hierarchical reinforcement learning". Discrete event dynamic systems 13.4 (2003): 341-379.
14. S Pateria., et al. "Hierarchical reinforcement learning: A comprehensive survey". ACM Computing Surveys (CSUR) 54.5 (2022): 1-35.
15. JM Steigerwald, SP Murarka and RJ Gutmann. Chemical mechanical planarization of microelectronic materials. John Wiley & Sons

(1997).

16. H Lee, H Jeong and D Dornfeld. "Semi-empirical material removal rate distribution model for SiO2 chemical mechanical polishing (CMP) processes". Precision Engineering 37.2 (2013): 483-490.
17. BE Stine., et al. "Rapid characterization and modeling of pattern-dependent variation in chemical-mechanical polishing". IEEE Transactions on Semiconductor manufacturing 11.1 (1998): 129-140.
18. D Ouma., et al. "An integrated characterization and modeling methodology for CMP dielectric planarization". in Proceedings of the IEEE 1998 International Interconnect Technology Conference (Cat. No. 98EX102). IEEE (1998): 67-69.
19. A Maury., et al. "A modification to Preston's equation and impact on pattern density effect modeling". Advance metallization and interconnect system for ULSI application (1997).