Srinivasa Rao Gondi
Sr. Principal test engineer, NXP semiconductors San Jose, United States

Received: 14-08-2023; Revised: 30-09-2023; Accepted: 28-10-2023

Adaptive Test Optimization for Reducing Cost and Improving Coverage in High-Volume Semiconductor Production

Abstract
High-volume semiconductor production requires fast and reliable testing to control cost while protecting product quality. Conventional full testing provides strong fault coverage, but it increases test time, test cost, and production burden when every device follows the same fixed test flow. Adaptive testing offers a better direction by selecting test paths according to device response, wafer-level risk, and coverage importance. Existing test optimization methods often reduce time or cost, but they may not fully control escaped defects when test removal is too aggressive. This article presents an adaptive test optimization framework for reducing test cost and improving coverage in semiconductor production. The study compares conventional full testing, rule-based adaptive testing, machine-learning-assisted adaptive testing, and aggressive test reduction using simulated high-volume production data. The results show that machine-learning-assisted adaptive testing gives the best balance by reducing test time and cost while maintaining high fault coverage, improving screening accuracy, and lowering escaped defects. The framework supports safer and more efficient semiconductor test decision-making.
Keywords: adaptive testing; semiconductor production; test optimization; fault coverage; escaped defects.

1. Introduction

High-volume semiconductor production depends on fast, accurate, and cost-effective testing to ensure that only reliable devices move to packaging and shipment. Test operations are necessary because small process shifts can create hidden electrical failures, weak devices, or early-life reliability risks. Adaptive testing can reduce unnecessary test steps by using production data, device response patterns, and decision rules to select the most useful tests for each wafer or device group. Low-cost adaptive testing can support high-yield integrated circuit production by reducing test burden while maintaining screening quality [1]. This makes adaptive test optimization important for manufacturers that need both high throughput and strong quality control.

Semiconductor test research has moved from fixed test flows toward data-driven test decisions that respond to product and wafer behavior. Wafer-level adaptive testing can use collaborative decision models to decide whether additional tests are needed for specific wafer regions or product groups [2]. Parameter correlation analysis can reduce test escapes by identifying weak devices that may pass simple screening but fail under broader quality checks [3]. Machine learning methods can detect wafer-test-induced defects and improve screening sensitivity in production testing [4]. Repeated test schemes can improve integrated circuit test quality and support near-zero-defect targets [5]. These studies show that adaptive testing must balance cost reduction, fault coverage, and escaped-defect control.

The main limitation in existing work is that many test optimization methods focus on only one target, such as reducing test time, improving screening accuracy, or lowering test escapes. This separated approach can create practical problems because aggressive test reduction may increase defect escape, while excessive retesting can increase cost and reduce throughput. Multi-correlation screening can improve escape detection by examining relationships among test parameters [3]. Repeated test schemes can improve quality, but they may increase total test time when they are not guided by adaptive decision logic [5]. The core research problem addressed in this study is the limited ability of existing adaptive test methods to jointly reduce test cost, preserve fault coverage, and control escaped defects under high-volume semiconductor production conditions.

This problem matters because semiconductor manufacturers must control test cost without weakening product quality or customer reliability. A useful adaptive test strategy should identify which devices need full testing, which devices can follow a reduced test path, and which devices require additional screening because of abnormal response patterns. The conceptual direction of this article is to develop an adaptive test optimization framework that combines test-response data, risk-based decision rules, model-assisted screening, and coverage-aware test selection. The key contribution of this study is a risk-based adaptive test-path selection approach that reduces unnecessary testing while protecting fault coverage and limiting escaped defects in high-volume semiconductor production.

2. Methodology

This study used a simulation-based adaptive test optimization design for high-volume semiconductor production. The simulated production dataset represented 120 fabrication lots, with 4,000 devices per lot and 480,000 total device records. Each record included wafer-level test data, die-level electrical responses, bin outcomes, retest indicators, wafer position, and final quality labels. K-nearest neighbor classification has been applied to adaptive VLSI test cost reduction by identifying devices that may not require the same full test path [6]. This study followed the same general direction but used a broader decision structure that considered cost, coverage, escaped-defect risk, and screening confidence together. The overall design was intended to reflect realistic production testing where every saved test second can affect total manufacturing cost.

The dataset was arranged to compare three testing strategies: conventional full testing, rule-based adaptive testing, and machine-learning-assisted adaptive testing. Conventional full testing applied the complete test sequence to all devices without considering device risk. Rule-based adaptive testing skipped selected low-value tests when the device response remained safely inside the pass region. Machine learning classification can reduce VLSI test cost by separating useful and less useful test patterns from production response data [7]. The machine-learning-assisted strategy used device response patterns, wafer-neighborhood behavior, and coverage importance to decide whether a device should receive reduced testing, full testing, or additional screening. This comparison design made it possible to measure both economic benefit and screening quality.

The adaptive optimization model used test variables and decision parameters that reflected practical semiconductor test conditions. Improved classification-based VLSI testing can reduce test burden when the model separates informative test items from low-value test items with acceptable reliability [8]. Table 1 presents the test optimization variables and decision parameters used in the adaptive semiconductor production testing framework. The variables covered test time, test cost, parametric weakness, retest stability, wafer-level spatial risk, and fault-coverage importance. These inputs helped the system decide whether test reduction was safe or whether additional screening was needed. The table also shows how each variable supported cost reduction, risk control, or coverage protection.

Table 1. Test optimization variables and decision parameters used in adaptive semiconductor production testing

Variable / parameter	Data source	Measurement form	Optimization purpose	Adaptive test action
Test execution time	Automated test equipment log	Seconds per device	Reduce avoidable test time	Skip low-value tests for low-risk devices
Test cost weight	Production cost record	Cost per test item	Improve cost efficiency	Prioritize costly tests for optimization

Parametric fail margin	Electrical test response	Distance from specification limit	Protect screening sensitivity	Retain or extend tests for marginal devices
Retest agreement score	Repeat-test record	Pass/fail consistency	Reduce false pass decisions	Add retest for inconsistent outcomes
Wafer neighborhood risk	Wafer-map data	Local risk score	Capture local failure patterns	Increase screening near risky regions
Coverage criticality score	Fault-model database	Coverage importance level	Preserve fault coverage	Do not remove essential tests

The preprocessing stage converted raw test records into clean and comparable model inputs. Missing test values were separated into skipped-test values and measurement-failure values because these two cases have different meanings in production testing. Valid test pattern identification can support adaptive VLSI testing when machine learning separates informative patterns from redundant patterns [9]. Electrical measurements were normalized so that voltage, current, leakage, and timing values could be used within the same model. Extreme values near specification limits were retained because they may represent weak devices rather than simple noise. This preprocessing method reduced measurement inconsistency while protecting useful failure information.

Feature engineering was used to build risk indicators from raw test responses. The fail-margin index measured how close a device was to the pass/fail boundary, while the local wafer-risk score used nearby die behavior to detect spatially clustered risk. The test redundancy score estimated whether one test item repeated information already captured by another test item. The retest instability score measured disagreement between first-pass and repeat-test outcomes. These features allowed the adaptive system to make decisions using both individual device behavior and wider wafer-level production patterns. The final feature set was designed to support safe test reduction instead of blind test removal.

The adaptive test decision model assigned each device to a reduced test path, full test path, or extended screening path. Low-risk devices entered the reduced test path when the fail-margin index was safely above the acceptance band, wafer-neighborhood risk was low, and removed tests had low coverage criticality. High-risk devices entered the extended screening path when their fail margins were narrow, retest agreement was weak, or nearby wafer regions showed abnormal failure behavior. Convolutional-neural-network-based adaptive testing can use production response patterns to support test-flow decisions when spatial or response-based information is available [10]. Medium-risk devices followed a controlled full test path in which only non-critical redundant tests were considered for removal. This logic allowed the method to reduce test cost while still protecting important fault-detection capability.

The proposed framework was evaluated by comparing conventional full testing, rule-based adaptive testing, and machine-learning-assisted adaptive testing under the same simulated production conditions. The main evaluation measures were test time reduction, test cost reduction, fault coverage, screening

accuracy, escaped-defect rate, retest load, and production throughput. The adaptive strategy was accepted only when test cost decreased without a meaningful loss of fault coverage or an increase in escaped defects. A 70:15:15 training, validation, and testing split was used at the lot level to reduce data leakage between related devices. Model performance was also checked across low-risk and high-risk wafer regions to confirm that the method did not perform well only on easy production cases. This evaluation design made the framework suitable for high-volume semiconductor testing, where cost efficiency and product quality must improve together.

3. Results and Discussion

The simulated production results showed that adaptive test optimization reduced test burden while maintaining strong screening quality. Conventional full testing provided the highest fault coverage, but it also required the longest test time and the highest cost per device. Table 2 compares conventional full testing, rule-based adaptive testing, machine-learning-assisted adaptive testing, and aggressive test reduction under the same production conditions. The machine-learning-assisted strategy gave the best balance by reducing test time from 12.8 s to 8.1 s per device while maintaining 98.6% fault coverage. It also reduced cost per device from USD 0.184 to USD 0.119 and improved throughput from 281 to 444 devices per hour. This result shows that adaptive testing can improve production efficiency when test-path decisions are guided by risk, coverage value, and device response quality.

Table 2. Comparative performance of conventional and adaptive test optimization strategies

Test strategy	Test time per device (s)	Test cost per device (USD)	Fault coverage (%)	Screening accuracy (%)	Escaped defects (ppm)	Throughput (devices/hour)
Conventional full testing	12.8	0.184	99.1	97.8	38	281
Rule-based adaptive testing	9.4	0.137	97.9	96.9	52	383
ML-assisted adaptive testing	8.1	0.119	98.6	98.3	31	444
Aggressive test reduction	6.9	0.101	95.8	94.7	86	522

Figure 1 shows the single-panel comparison of test cost, test time, and fault coverage under adaptive optimization. The ML-assisted adaptive strategy reduced both cost and test time more effectively than rule-based adaptive testing while keeping fault coverage close to the conventional full-test baseline. This outcome occurred because the ML-assisted method did not remove tests only to shorten the test flow. It used fail-margin behavior, wafer-neighborhood risk, and coverage criticality to decide which tests could be safely skipped. Aggressive test reduction achieved the lowest cost and shortest test time, but the drop in fault coverage shows that uncontrolled test removal is not suitable for reliable production

testing. Figure 1 therefore confirms that adaptive optimization must protect coverage while reducing unnecessary test effort.

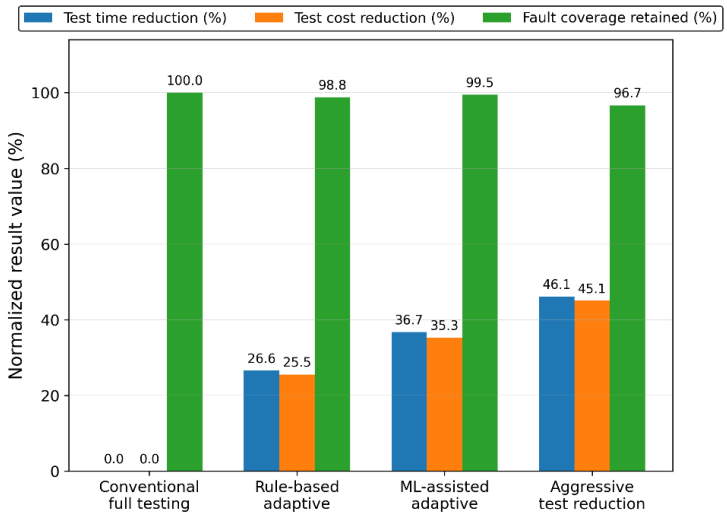

Figure 1. Test cost, time, and coverage under adaptive optimization

Figure 2 shows screening accuracy, escaped-defect rate, and throughput across the same test strategies. The ML-assisted adaptive strategy improved screening accuracy to 98.3% and reduced escaped defects to 31 ppm while increasing throughput to 444 devices per hour. This result is important because it shows that the proposed strategy improved speed without weakening defect control. The aggressive reduction strategy gave the highest throughput at 522 devices per hour, but it increased escaped defects to 86 ppm, which is not acceptable for quality-sensitive semiconductor production. Rule-based adaptive testing improved time and cost, but its escaped-defect rate was higher than the ML-assisted method because it relied on fixed decision limits. Figure 2 shows that the best test strategy is not the fastest one, but the one that improves throughput while protecting screening reliability.

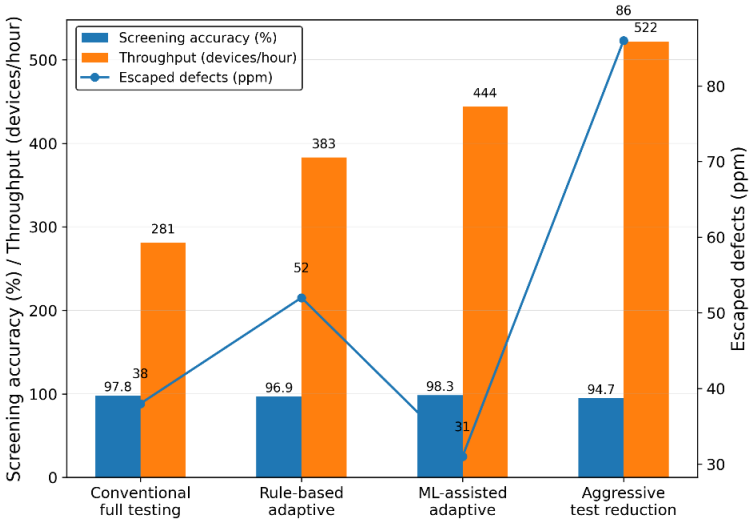

Figure 2. Screening accuracy, defect escape, and throughput across test strategies

The overall results show that ML-assisted adaptive testing gave the strongest practical outcome for high-volume semiconductor production. Conventional full testing remained reliable for fault coverage, but it did not provide cost or throughput advantages. Rule-based adaptive testing reduced test burden, but it was less effective in controlling escaped defects because it could not fully adapt to device-level and wafer-level risk patterns. Aggressive test reduction improved test speed, but its high escaped-defect rate showed that cost reduction alone is not a safe optimization target. The proposed ML-assisted strategy performed better because it combined device response, fail-margin information, wafer-neighborhood risk, and coverage criticality in one adaptive decision process. These findings confirm that adaptive test-path selection can reduce cost, improve throughput, and preserve production quality when the decision logic is guided by both risk and coverage importance.

4. Conclusion

This study presented an adaptive test optimization approach for reducing test cost and improving coverage in high-volume semiconductor production. The proposed method compared conventional full testing, rule-based adaptive testing, machine-learning-assisted adaptive testing, and aggressive test reduction under the same simulated production conditions. The results showed that simple test removal can reduce time and cost, but it can also increase the risk of escaped defects. This confirms that semiconductor test optimization must balance speed, cost, coverage, and product quality together. The findings also show that test-flow decisions should be guided by device risk rather than by cost reduction alone. This makes adaptive testing more suitable for high-volume production environments where quality and throughput must be maintained at the same time.

The machine-learning-assisted adaptive strategy gave the best overall performance in the simulated production setting. It reduced test time and test cost while keeping fault coverage close to the conventional full-test baseline. It also improved screening accuracy and reduced escaped defects compared with the rule-based adaptive method. This happened because the ML-assisted strategy used device response, fail-margin behavior, wafer-neighborhood risk, and coverage criticality together before selecting the test path. The method avoided the weakness of aggressive test reduction by retaining important tests for risky devices and removing only low-value tests for stable devices. This shows that intelligent test selection can improve production efficiency without weakening screening reliability.

The main contribution of this article is a practical adaptive test-path selection framework that supports safe test reduction in semiconductor production. The method can help manufacturers reduce unnecessary testing, improve throughput, and protect important fault-screening steps. It also provides a structured way to study the trade-off between test cost, fault coverage, throughput, and escaped-defect control in semiconductor production. Future work can extend this framework by using real production

datasets, adding more device families, and testing the approach across different technology nodes. Further validation with production-scale data would improve the reliability and industrial usefulness of the proposed method. This direction can make adaptive testing more practical and scientifically stronger for large-scale semiconductor manufacturing.

References

1. Pan, Y., Liang, H., Li, J., Huang, Z., Yi, M., & Lu, Y. (2025). Low test cost adaptive testing method for high yield IC products. *Integration*, *103*, 102401.
2. Pan, Y., Liang, H., Li, J., Qu, J., Huang, Z., Yi, M., & Lu, Y. (2024). Wafer-level adaptive testing based on dual-predictor collaborative decision. *Journal of Electronic Testing*, *40*(3), 405-415.
3. Zhang, J., You, H., Jia, R., & Wang, X. (2022). The research on screening method to reduce chip test escapes by using multi-correlation analysis of parameters. *IEEE Transactions on Semiconductor Manufacturing*, *35*(2), 266-271.
4. Cheng, K. C. C., Chen, L. L. Y., Li, J. W., Li, K. S. M., Tsai, N. C. Y., Wang, S. J., ... & Hsu, C. L. (2021). Machine learning-based detection method for wafer test induced defects. *IEEE Transactions on Semiconductor Manufacturing*, *34*(2), 161-167.
5. Yeh, C. H., & Chen, J. E. (2022). Application of three-repetition tests scheme to improve integrated circuits test quality to near-zero defect. *Sensors*, *22*(11), 4158.
6. Song, T., Huang, Z., Zhang, L., Hong, Q., Yang, Z., & Krstic, M. (2024). Test cost reduction for VLSI adaptive test with K-nearest neighbor classification algorithm. *IEEE Transactions on Circuits and Systems II: Express Briefs*, *71*(7), 3508-3512.
7. Song, T., Huang, Z., & Yan, A. (2022). Machine learning classification algorithm for VLSI test cost reduction. *Integration*, *87*, 40-48.
8. Huang, L., & Song, T. (2022). VLSI test through an improved LDA classification algorithm for test cost reduction. *Microelectronics Journal*, *125*, 105461.
9. Song, T., Huang, Z., & Guo, X. (2022). RLDA: Valid test pattern identification by machine learning classification method for VLSI test. *Microelectronics Journal*, *128*, 105549.
10. Shin, D., Oh, S., Kim, W., & Kim, H. (2025). A Novel Adaptive Testing Method Using Convolutional Neural Networks. *JOURNAL OF SEMICONDUCTOR TECHNOLOGY AND SCIENCE*, *25*(5), 610-615.