Frontiers in Emerging Engineering & Technologies (Open Access)

Volume 01, Issue 01, April 2024,
Publish Date: 04-16-2024
PageNo.49-73

Challenges in Verifying Test Pattern Generators for Modern Semiconductor Designs

Vikas Nagaraj
MTS at Advanced Micro Device (AMD), San Jose, California, USA

ABSTRACT

The Modern SoEs are heterogeneous IP, tester I/O constraints, dynamically power-gated regions, and X-prone interfaces, and verification of SoESe Tunable TPG, therefore, is a multi-objective process of controllability, observability, aliasing, and power integrity. This work characterizes TPG verification as generating traces that provide a demonstration that patterns satisfy quantifiable targets below realistic clocks/resets/constraints/tester limits, and under five different kinds of TPGs: deterministic ATPG, LBIST (rereading/weighting), and compression-based TPGs. The approach integrates static DFT lint and formal BER properties, X-aware simulation and fault simulation, emulation, statistical confirmation (bootstrap confidence and MISR alias analysis), and power-aware vet and veto/scheduling. At three technology nodes, it has been observed that 96-channel otherwise-good distribution maintains ~99% stuck-at and ~94% transition coverage, reducing the pattern volume by ~38%; reseeded LBIST with 48-bit MISR may improve transition coverage and time; and selective bitwise X-masking may recover observability with minimal loss; and empirical results concerning alias behavior support theoretical alias behavior, and n=48 seems to be effective on complex designs. It documents Pareto improvements to coverage, safe toggle envelopes, and a silicon-correlation plan that aligns the coverage-driven modeling projections with ATE signatures. The result is a defensible sign-off contract-coverage with confidence intervals, alias-ppm targets mapped to MSR width, toggle/IR limits, constraint legality, equivalence proofs, ATE fit, and machine-readable manifest-making DFT verification defensible evidence.

KEYWORDS: *ATPG (deterministic automatic test pattern generation), LBIST (LFSR stimulus / MISR response compaction), Test compression (decompressors, phase shifters, channel utilization), X-propagation control (X-masking / X-bounding for observability), MISR aliasing analysis.*

1. Introduction

Modern system-on-chip systems combine heterogeneous IP into multi-domain clocking and voltages, and they depend on retention cells, isolation cells, and intensive power gating. These facts increase controllability and observability during scan, introduce timing hazards on scan-enable and test clocks, and cause an increase in the frequency of unknown (X) values due to a lack of initialization state, as well as macro analogs and domain crossing. There is limited and costly tester I/O, driving the use of compressed scan where a limited number of ATE channels are connected to on-chip decompressors and phase shifters. Safety-critical applications like automotive, where ISO 26262 ASIL targets are needed, provide evidence of a non-violation of functional safety properties and diagnostic coverage at the target level. In the meantime,

scaling technology extends the defect continuum beyond shorts [10]. It opens to timing, resistance bridges, and defects internal to the cell, all worsening the need to make reliable pattern generation and verification methods. The need to maintain business pressure on minimizing pattern volume and tester seconds creates additional pressure, as verification must demonstrate that compaction, masking, and power-aware scheduling do not reduce coverage or introduce quality risks hidden during verification. In practice, these forces result in TPG verification being a

multi-objective engineering challenge that ranges across logic design, DFT architecture, as well as power integrity and statistical assurance.

In this paper, verification of a test pattern generator refers to the production of justifiably traceable evidence that selected patterns are produced with respect to objectives that are measurable on realistic assumptions about clocks, resets, constraints, and tester limits. Targets include a full range of fault coverage across the applicable fault models--stuck-at, transition, path-delay, bridging, and cell-aware--without violating any mode or pin constraints; and an adequate reduction in aliasing risk, whether analytically by selecting the MISR width n to satisfy a theoretical upper bound near 1/2^n or empirically by a program of stressing the circuit and proving that aliasing risk has been reduced to reasonable levels. The scope includes deterministic ATPG, logic built-in self-test with reseeding or weighting, and compression-based TPGs that map tester channels to scan chains via phase-shifters and on-chip decompressors. Verification is done under tester mode constraint with legal states, scan-enable protocols, and safe clocks are modeled.

The research adds an X-aware and power-aware verification approach that will provide statistical confidence and a feasible path to rocking silicon. The procedure combines textual lint with formal property checks, ensuring that scan-enable exclusivity, muxes that are correct on test-mode crossings, decompressor and mask controller safety, and lock-up latch compatibility are achieved. Simulation of stuck-at, transition, and path-delay coverage under realistic constraints is configured to measure with stuck-at faults, transition faults, and path faults on a gate-level directed to X-propagating fault simulation. Statistical methods are used to calculate confidence bounds of achieved LBIST coverage, seed-to-seed variance in LBIST, and correlate the effects of MISR width and response correlation into an aliasing parts-per-million value. Power-wise vetting blocks sequences that use more than preset toggle/IR power budgets, and sequencing prevents dangerous bursts. The result is a sign-off checklist that becomes an executable asset and also a reusable artifact set including configuration files, seeds, property libraries, logs, and machine-readable reports that can be audited and replayed to inform quality gates and safety assessments.

This research is organized into chapters. Chapter 2 presents a survey on the existing work in ATPG, LBIST, compression architectures, and verification methods, and gaps in X-prop fidelity, alias validation, and power-aware sign-off. Chapter 3 describes datasets, data preprocessing, visual analytics, the generators to be verified, and the objective thresholds and metrics. The fourth chapter describes a framework of verification execution sequencing consisting of static, formal, simulation, emulation, and statistical, as well as automating CI/CD. Chapters 5-8 provide experiments, discussion, limitations, and conclusions.

2. Literature Review

2.1 Classical ATPG & Fault Models

Classical automatic test pattern generation (ATPG) associates physical defect drivers with logical effects that are controllable and observable using scan. Stuck-at models assume a line is forever forced to logic-0 or logic-1 and remains canonical because it scales to huge designs, correlates with a great deal of opens/shorts, and can act as a grounded reference point against which controllability and observability can be assessed. Transition faults model timing-based faults as slow-to-rise or slow-to-fall; therefore, generators need to support two-cycle-launch/capture semantics, domain-specific clocking, and on-chip variation. Path-delay faults have the effect of sensitizing only specific critical paths and thus raising the requirement of realistic constraints, launch modes, and false-path pruning. Bridging faults connect adjacent signals to rails; effective patterns demand orthogonal control of the bridged nets to induce wired-AND/OR behavior and enough observability at capture points.

The ATPG tool will read the design netlist, test procedure file, ATPG library, and fault list, and generate scan test patterns and ATPG info files as shown in the figure below [18]. The flow is focused on classical fault models, such as stuck-at (logic-0/1), transition (slow-to-rise/fall with two-cycle launch/capture and domain clocks), and path-delay (critical-path sensitization with realistic constraints and false-path pruning), as well as bridging (adjacent-net shorts requiring orthogonal control to provoke wired-AND/OR) with preservation of observability at capture points. This pipeline balances controllability and observability with the limitations of testers.

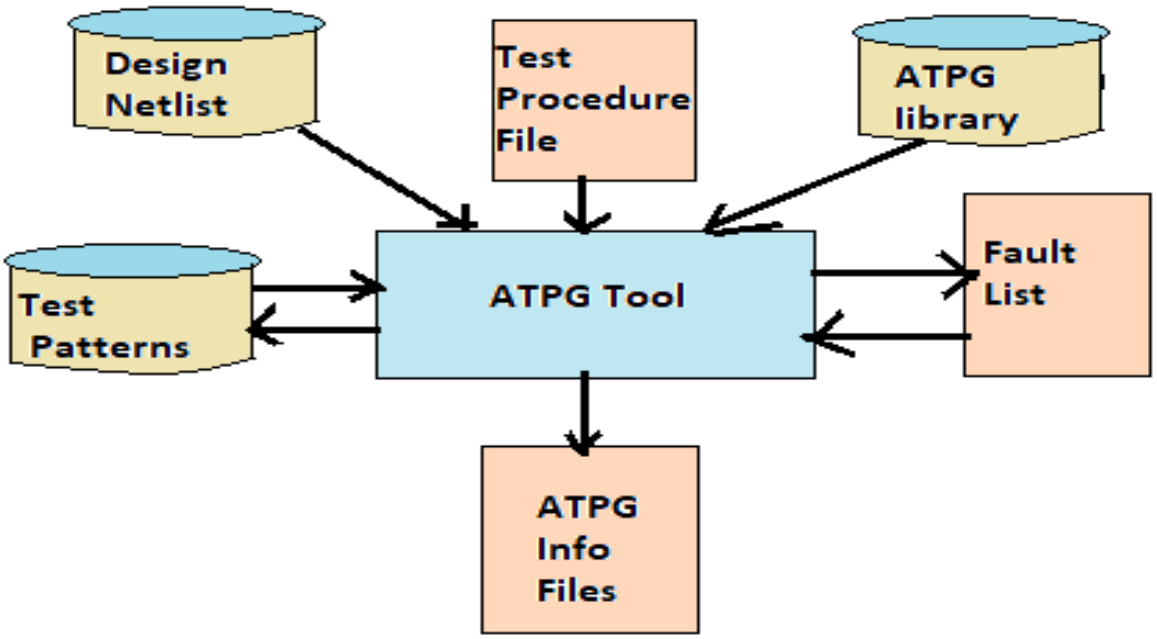

Figure 1: ATPG workflow: inputs, fault models, and generated scan test patterns

Cell-aware testing increases the level of granularity to the transistor level of defects; the patterns must reveal internal arcs and weak drives; this increases the controllability requirements and makes it harder to carry out the compaction. The models have tangible budgets in test-pattern generator (TPG) verification. Controllability is sensitive to decompressor degrees of freedom and suppressing cubes with constraints. Observability is sensitive to scan ordering, compaction style, and the presence of unknowns (X), introduced by decompression of uninitialized flops and retention cells, or decompression of mixed-signal wrappers or synchronous boundaries. Structural measures like the SCOAP score, hard-to-control and hard-to-observe cones, and maintaining verification are targeted at the real bottlenecks. More importantly, well-defined context boundaries are required to determine where fault activation/observation is required, either at MISR inputs, chain heads, or core wrapper pins, because vague boundaries conceal escapes and misrepresent coverage accounting. This lesson is also consistent with the need to define strong bounded contexts before breaking a complex system into smaller components [5].

2.2 Pseudo-Random/BIST Fundamentals

The built-in self-test (BIST) of Logic uses linear feedback shift registers (LFSRs) to create test stimulus and uses multiple-input signature registers (MISR) to compact results [1]. Primitive polynomials can give the maximal-length sequences, and a judicious use of phase-shift registers and the position of taps reduces the effects of aliasing between lines and spreads the switching between scan chains more uniformly. Probability biasing of the 0/1 probability through weighted generating random patterns supports the optimization of random-resistant structures such as long-AND cones or gated clocks, and reseeding inoculates deterministic states to access rare cubes without causing the explosion of external pattern volume. MISR compression maps long response streams to fixed-width signatures; in the ideal linear model, the upper bound on the alias probability scales asymptotically as $1/2^n$ in n-bits of output width, but in practice, the probability of aliasing depends on the correlation and X-masking, and the designers will be unaware of how often unknown blocking will occur.

The use of pseudo-random toggling in LBIST induces power integrity, as functionality exercised in LBIST can be higher than in actual use; capture windows must adhere to toggle envelopes in domains, clock dividers, and power-gating. Common mitigations are launch-rate throttling, bias schedules, and chain staggering to lessen concurrent switching. Verificationally, LBIST quality is as much brought about by orchestration as it is by algebra. Modeling configuration as a gated pipeline, seed generation, reseed packing, MISR polynomial selection, and signature thresholds allow automated go/no-go checks: an upper bound on the minimum incremental coverage of each reseed pack, a maximum capture-window toggle density, and an upper bound on alias-risk. Such an attitude resembles the way continuous

integration pipelines work, where promotion gates are represented by static and dynamic analyzers, focusing on automation, repeatability, and quality rules that the developer has established before an artifact can be promoted to further stages [15].

2.3 Test Compression Architectures

Compressed scan minimizes both external channel count and tester memory, compared with the internal controllability. Linear decompressors multiply a few incoming channels into many internal chains with a statistical XOR network or rearrange an incoming chain with a structured XOR network and optional phase shifting [21]. Response compaction, Space, time, or both, reduces the output to manageable signature volumes, and dynamic X-masking or X-bounding logic prevents unknowns from corrupting signatures. Since decompressors with many chains hook up to a small number of control bits, rippling is very significant; inefficiently-expressed constraints kill viable cubes and leave hole states where ATPG cannot find assignments. Verification must thus ensure that decompressor rank is adequate, channel-to-chain mapping does not approach pathoscopic coupling, mask controllers' legal timing (no enable allowed during shift capture overlap), and unbounded transformers exist within a domain of knowledge, not simulator faith.

Controllability can be measured by solving cube feasibility within channel limits; observability can be measured by tracing X propagation into the compactor, placing temporary observation points, or by running paired optimistic/pessimistic X-prop simulations. Practical delivery is not only a matter of schedule, but also of wiring. Channel distribution, mask-enable timing, and capture sequencing must be coordinated to address tester bandwidth, pin targets, and power/IR constraints. Practical schedules combine high-value bursts and intersperse cool-down periods or low-toggle interspersions to remain within thermal limits. The DARPA-sized decision space can be compared to notification planning in safety-critical operations: limited time windows to signal the device must be scheduled where they achieve the most benefit, subject to strict constraints, and timing and resource-aware planning can make a tangible difference to the quality of the outcomes [31].

2.4 Verification Techniques in DFT

Credible sign-off stacks add to other techniques. Static DFT linting finds broken chain DFTs/IL/SD/MD, illegal cross-asynchronous stitching, missing/mis-polarity lock-off latches, inverted/gated scan-enable, and clock-mux Misuse; also checking that test clocks are separate from functional PLLs and test mode turns off analog side condition effects like charge pumps. Formal property checking demonstrates semantic safety and reachability: scan-enable exclusive with shift, decompressor programming exclusive with capture, liveness of seed distribution over all phase shifters, and protections against test clocks interfering with test clock gating/alleviation. The approach of equivalence checking assures that insertion of DFT does not alter functional behavior under test mode disabled conditions, precluding accidental logic perturbation. Dynamic techniques offer evidence that can be executed.

DFT verification will begin at RTL and repeat static sign-off at logic/DFT synthesis and place-and-route stages before ATPG, as shown in the figure below. Lint checking detects broken chains, cross-domain stitching violations, lock-up polarity errors, as well as gatescanned/inverted scan-enable and clock-mux maladaptation. Formal properties ensure mode exclusivity of scan-mode, secure encoder programming, liveness of seed-distribution, and test clock non-interference. Equivalence check verifies that the insertion of DFT did not alter the functional behavior when the test is disabled, giving stage-gated evidence that this flow would be followed through to the executable results.

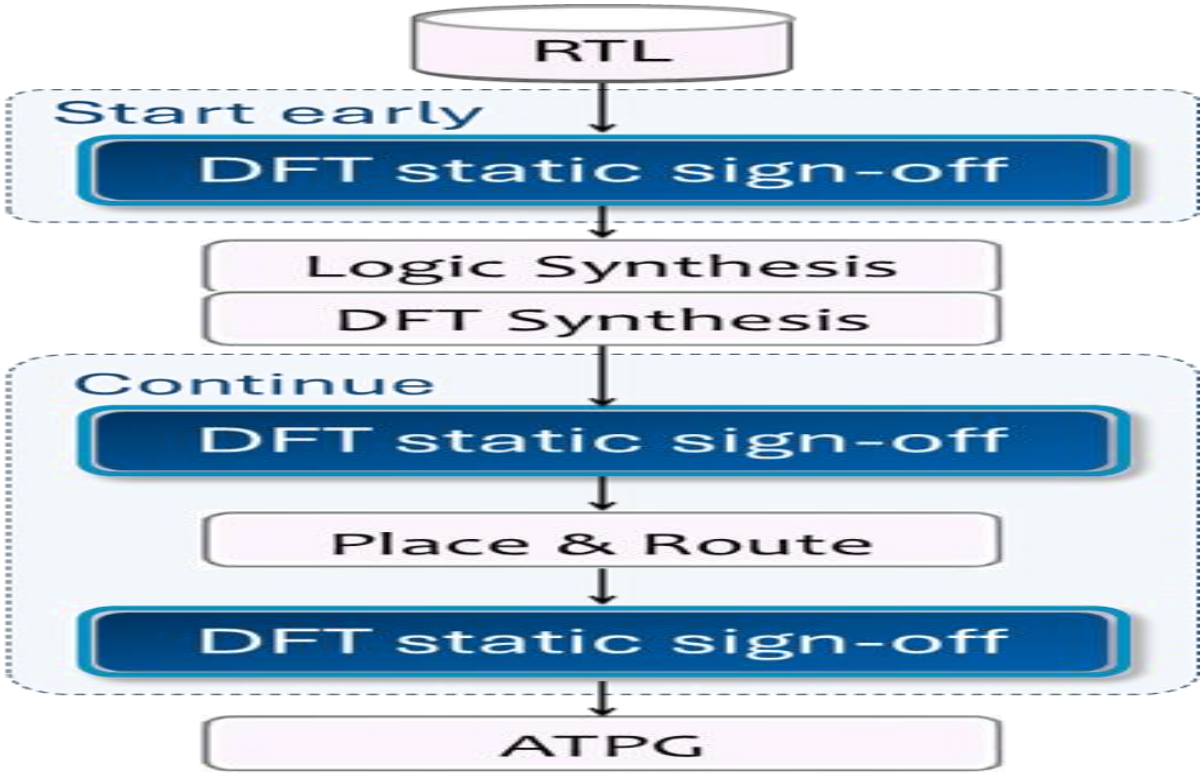

Figure 2: Iterative DFT static sign-off from RTL to ATPG

The simulation at the gate level demonstrates that the test modes, clock gating, isolation cells, and retention are as desired, and pairing optimistic and pessimistic X-prop settings bounds the uncertainty due to unknowns. By fault simulation, such as parallel, deductive, or concurrent fault simulation, one can estimate coverage efficiently and assign misses to causes of controllability or observability. Override power simulation. Power-aware simulation will overlay the toggle-based estimators or vector-wise rail analysis to ensure that capture events remain within the IR-drop envelopes. FPGA prototyping or emulation can provide MHz time scale execution, allowing the study of seed-to-seed variance, stress within fault families, and generation of intermediate signature sets that can be used to make a differential diagnosis. Automated dashboards and formalized feedback loops will turn these results into action: coverage state deltas, alias-risk tallies, and constraint violations each feed back to DFT owners in a worldwide cycle of continuous improvement that resembles how formative feedback systems operate with learners, shepherding them through optimized interventions [13].

2.5 Gaps & Industrial Challenges

Despite the depth of methodology, several production pain points remain. The recognized X-fidelity is disoriented and optimistic metrics overestimate coverage, whereas pessimistic metrics underestimate accurate detections; neither perfectly aligns with silicon given unknown voltages travel over reconvergent fan-out, mixed-signal substrates, or retention cells. A practical practice would be to surround the risk with the paired simulations, verify with the emulation signatures, and robustify/reinforce weak cones through selective mask placements or observation points. Power realism is brutal: pseudo-random capture may engage local hotspots that cannot be noticed with average toggle bits; vector-sensitive rail analysis and thermal guard-bands must reject unsafe spectral windows, at the counter-cost of pattern number.

Aliasing assurance can be brought down to quoting the $1/2^n$ bound, but correlated responses, non-linearities in compaction, and obscurations between colliding faults complicate matters; empirical collision studies by fault family are needed. Management of constraints is brittle, and possible combinations of legal values include safety, package straps, and multiplexing of testers; partial ones consume ATPG resources on impossible cubes or fail retention/isolation rules. Reproducibility needs to be addressed as such a commodity: seed/version pinning, machine-readable reports, and artifact hashing are all necessary to enforce trustworthy sign-off and post-silicon correlation if escapes still occur.

3. Methods and Techniques

3.1 Data Sets and Design Portfolio

The verification campaign used production-level-scaled netlists and public benchmarks, as part of a mix of benchmarks, to exercise the test pattern generators (TPGs) over a realistic structural mix—bring-up Proceeds with Scan-based circuits. ISCAS and ITC were used during the bring-up of scan-based circuits as well because their documented fault lists and statistics allow quick verification of fault simulation, scan stitching, and reporting pipelines [34]. In order to resemble modern systems-on-chip (SoCs), industrial gate-level netlists were ingested in a nondisclosure agreement. These designs ranged in scale, from hundreds of thousands to tens of millions of flip-flops, with scan chains between a few dozen and several thousand, and possibly heterogeneous power and voltage islands. All netlists included scan-insertion reports, timing abstractions, and, when provided, power-grid views to check IR drops under test.

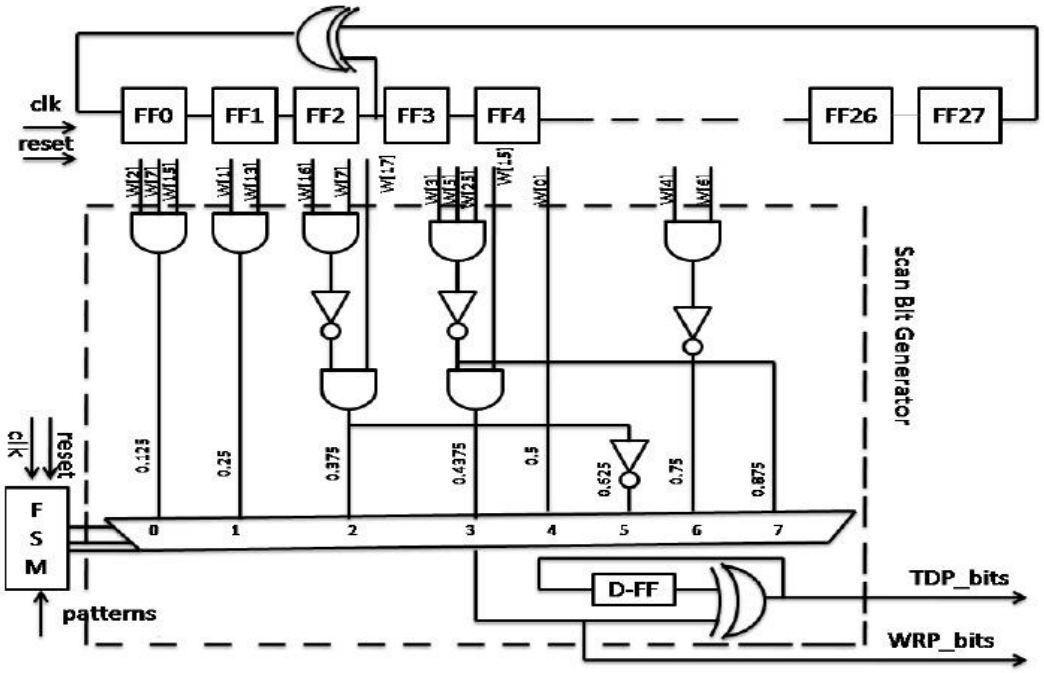

Figure 3: Scan-based test pattern generator block for benchmark and industrial netlists

A scan-based test pattern generator (TPG) and scan-bit generator structure were tested against our datasets, including ISCAS/ITC benchmarks, as well as production gate-level netlists under NDA, as shown in the figure below. The configuration is fitted to scan stitching, fault-simulation pipeline, and touch reporting and scaling up to tens of millions of flip-flops and up to thousands of scan chains. A design can include heterogeneous power/voltage islands. Related artifacts-a scan-insertion report and timing abstracts are generated, and optional power-grid visualizations can be generated to enable IR-drop tests both during test bring-up and volume checking at scale.

To facilitate a valid comparison whilst maintaining confidentiality, the identities of designs were substituted with parameters describing designs. The balancing quality of the chains was calculated as a ratio between the maximum and minimum chains; ratios higher than 1.5 prompted a recommendation to rebalance to decrease the shift-time variance and peak scan current. A measure of clocking complexity was the number of asynchronous domain pairs in the inter-domain scan connections; greater values necessitated more lock-up latches to defend against hold hazards at shift. Power topology descriptors provided island membership per scan cell, isolation and retention plans, as well as determination of the presence of a clock-stopped domain in test mode. As each design was a registered context, using contracts to specify functional requirements on each of the clocks, resets, legal primary-input settings, and power states, minimized configuration interleakage and enhanced reproducibility in interchanging projects [6].

Fault lists were created daily on a design and fault model basis. In stuck-at (SA), structural dominance and equivalence were used to collapse the set, with representatives left to contain pointers to dominated sites of diagnostic remapping. TF targets were generated with sign-off timing that excluded false and multi-cycle paths; launch/capture domain-specific constraints were introduced in the ATPG environment. Path-delay (PD) terminals were tabulated in timing abstracts and filtered

along a depth and slack dimension [11]. Fault information included metadata - topological location, pertinent cells in the scan, controlling resets and clocks, and power-domain membership - so that subsequent analyses could stratify the coverage by structure and operating condition. Labels reflected detected/undetected states, X-contamination (unidentified alarms on the path of propagation), and alias-suspect signatures when compacted by a multiple-input signature register (MISR).

3.2 Pre-processing & Feature Extraction

Tool-dependent variation was eliminated by normalizing the Netlists. The flow integrated standard-cell libraries, normalized scan-enable polarity, and included explicit tie-offs, and only needed to flatten the boundaries needed to expose scan stitching while maintaining black-boxes of analog PHYs, eFuses, and OTP. The trees were synchronous or asynchronous, and asynchronous reset trees that crossed domains were paired with lock-up latches to satisfy the shift-mode holds. Scan mode clock crossings were counted and synchronizers marked in such a way that at-speed campaigns would not have to make insecure assumptions about waveforms. A DFT lint pass ascertained the chain connections, scan-mux assurance, the controllable and observable gate log, and the mux exclusivity in test mode before any ATPG or LBIST was run.

Constraints were consumed in only one source of truth. Any legal primary-input settings, disablement of PLLs, test clock ratio setting, and mode-mux choice were presented in machine-readable sets. Powering limitations were maximum toggles per N-cycle window, and average toggles per domain limits based on sign-off IR-drop budgets. Tester constraints, such as vector memory, channel count, pin map, and maximum scan frequency, were programmed so that TPG optimization considered physical deployment [20]. It allowed close relationships between constraints (a disabling of a root clock would require disabling of its derived clocks unless overridden explicitly) and per-campaign manifests to allow runs to be reproducible.

Structure features were calculated to aid the diagnosis and to calculate the objectives. SCOAP-style controllability and observability indices were computed on the scan-configured netlist to measure justification and propagation effort; these were memoized per node and per domain. Interface topology features used at the graph-level were logic depth to the nearest observed point, fan-in/fanout, reconvergence degree, and distance to assigned scan cells of decompressor channels. In the cases where the design has compression (line insertion), a bipartite map between decompressor outputs and scan cells was used to estimate expansion fanout and detect channel starvation. LFSR polynomial taps, seed Hamming weight and linear complexity, phase-shifter topology, and MISR width were described by LBIST descriptors. All features were exported with the coverage results to allow later analyses to relate structural difficulty to a detection outcome and guide target setting at the block and domain level.

3.3 TPGs under Verification

Timing/power-aware dynamic compaction was used when deterministic ATPG was run. Sweeps were used to examine the trade-off of the patterns when X-fill, biased-fill, and minimum-transition fill policies were implemented. Explicit timing exceptions were imported in the speed-transition and path-delay of movements, and at-speed capture clocks were accurately phase-related to avoid fictitious slack. The generator produced per-pattern metadata: intended faults, scan-cycle activity profiles, required mode pins, scanned-chain testing activity profiles, and predicted switching envelopes. Any patterns that caused violations of the instantaneous or the sliding window toggle budget were vetoed or scheduled for spaced execution to stay within IR and thermal budgets.

The dynamic compaction flow (as shown in Figure 4) is integrated with deterministic ATPG using a verification flow. In parameter sweeps, the X-fill, biased-fill, and minimum-transition policies are compared as timing exceptions are imported at both the transition and path-delay tests. Phase-alignment of at-speed capture clocks avoids fictitious slack. The generator provides per-pattern metadata-target faults, scan-cycle activity, necessary mode pins, chain activity profiles, and predicted switching envelopes. Worst-case violating an instantaneous or sliding window toggle budget is vetoed or rescheduled, thereby ensuring IR-drop and thermal limits are not exceeded under testing conditions.

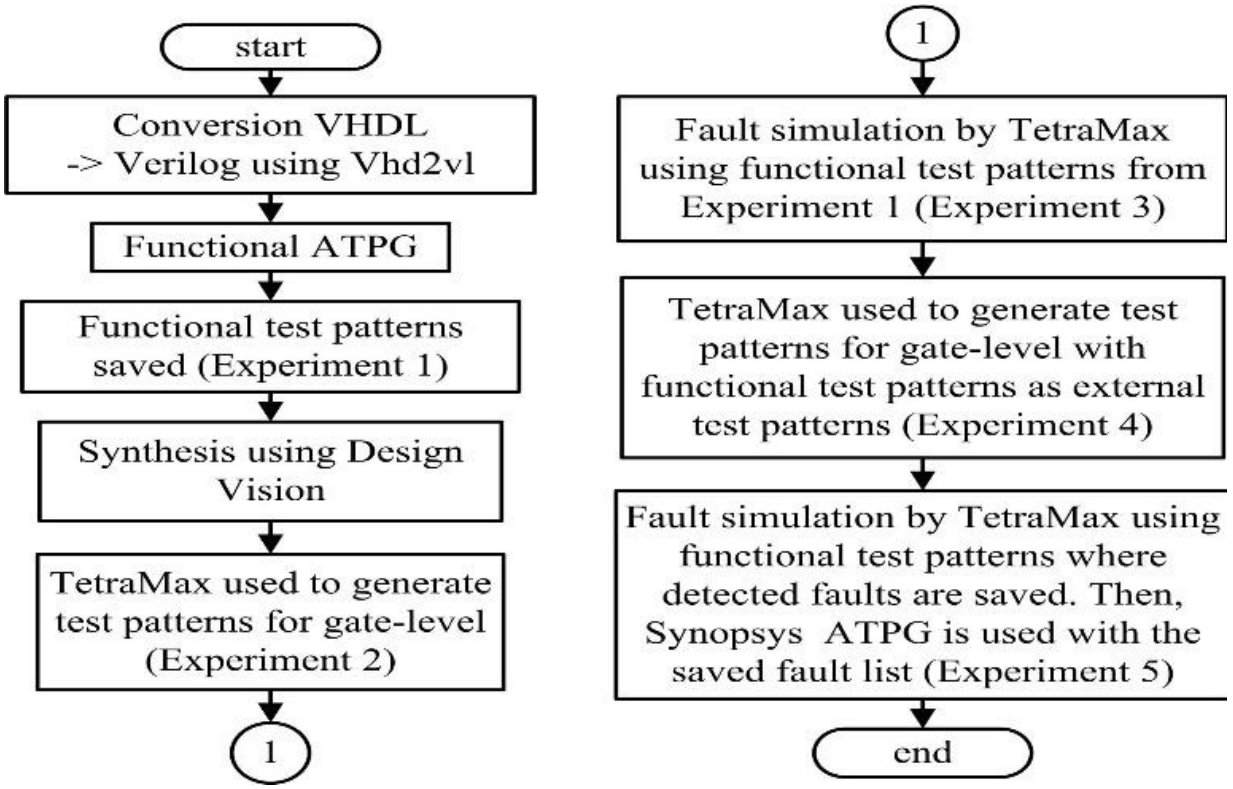

Figure 4: ATPG–fault simulation workflow with timing/power-aware pattern generation

The focus of verification with LBIST was seed cadence, phase-shifter strength, and MISR sizing. The cost-effective control seeding frequency was optimized to the levels of diminishing returns in terms of coverage, and at the same time, remained within the testers' time quotas. Phase shifters were modeled and measured in terms of decorrelation strength with pathologies where shared XOR networks unwittingly created correlation among scan inputs. MISR width n was chosen against an aliasing target, where the theoretical passthrough window (alias probability of roughly 2^{-n}) was taken as a target and then charged with correlated response families. Seed sets were designed to be as maximally linearly independent and spectrally diverse as clocking and power constraints allow. On-chip X-masking and X-bounding networks were checked to ensure their legality and to quantify the observability loss.

TPGs that use compression were modelled as explicit linear systems, and therefore expansion characteristics could be traced back to scan cells [26]. Channel utilization was monitored to indicate starvation at individual chains, and this could be related to faults unrecognized in deep reconvergent cones that were long-lasting. X-source-driven mask-controller behavior was recorded in terms of X-sources, including uninitialized SRAMs, black-box analog components, and clock-stopped islands; unknown blocking counts were computed on a chain-per-chain and domain-per-domain basis. Where recurring masking of certain logic cones was detected, the flow suggested additional observation locations or mask repacking. The equivalence and property-checking models were also updated with all compression paths and mask rules, and decompressor-to-scan connectivity to ensure consistency between dynamic simulation results and static guarantees.

3.4 Verification Objectives & Metrics

Goals struck a balance between structural purity, physical safety, and the costs of deployment. Fault model made coverage targets and to be in power/clock domains, with high-speed areas possibly lagging behind low-speed cells bearing acceptable timing margins. The compaction ratio was calculated as the ratio between unconstrained, naive pattern count and the resulting count after the application of dynamic compaction; sustained ratios greater than five exemplified successful compaction and posed well constraints. Pattern volumes were translated to tester time based on chain length distributions, scan frequency limits, capture windows, and vector memory ceilings to enable the benefits to be expressed in terms that the operations department understood, in terms of cost and throughput.

The risk of aliasing was simulated on two levels. A theoretical tier calculated the 2-subtracted bound based on MISR width under independence, and an empirical tier

estimated collision frequency via seed injection/phase-shifter configuration/ finer seed and phase-shifter scans. The more pessimistic of the two levels was chosen, and a confidence interval was added using bootstrap resampling across seeds and pattern windows. The difference in the maximum and minimum percent masked in a chain was also used to affect the alias estimate, and all such chains added less information to the signature and therefore increased the effective alias risk: this effect has also been quantified by replaying the mask during the compaction model.

Pre-silicon metrics were translated into outgoing quality expectations using the escape-rate proxies. RUD, based on scale, functional criticality, and usage rate of the affected logic, provides a structural proxy. A second proxy linked power integrity to detection reliability: the pattern-by-pattern current was convolved with cell-current models to estimate IR drop and thermal swings; patterns triggering IRs beyond safety limits were vetoed or spread in time with idle cycles to keep the capture within safety margins. To ensure decisions could be audited and could be repeated, rulers and measures were built into a centralized, machine-readable sign-off manifest, and an executive dashboard could track trend lines and regressions in the same way robust delivery pipelines track quality gates [16]. Excluded requirements were minimum coverage/model/domain with statistical confidence, maximum alias parts-per-million per MISR setting, upper limits on masked-logic percentage, pattern volume, and tester time/upper limits on power-on power envelopes. The data produced by each gate included the measured value, the threshold, and the method of measurement, and the links to evidence artifacts such as fault-simulation log, alias-stress matrices, and power-analysis reports, with a go/no-go decision being made safely at tape-out.

4. Verification Execution Framework for TPGs

4.1 Static DFT & Formal Property Checks

A stringent verification process for test pattern generators (TPG) starts with structural sanitation/property proofing at a demigated level. Scan-chain integrity Verification of continuity, ordering, and the absence of shorts or loops; failure makes observability invalid and noise not covered [3]. Lock-up latches are placed on every crossing between asynchronous or skewed clock domains to absorb shift-mode skew and to prevent hold failures as scan data ripples through chains. The structure also demands that scan-enable be single-source/glitch-free, mutually exclusive to functional enable, and gated such that during mission mode, the scan clocks cannot propagate. Input and output ports are mapped bijectively to decompressor and compactor outputs, and chain lengths are balanced to fit within tester time budgets. A property library that can be shared across an IP and parameterised by chain count, decompressor width, and mask topology adapts to the growth of an IP and avoids divergence between block-level and top-level properties.

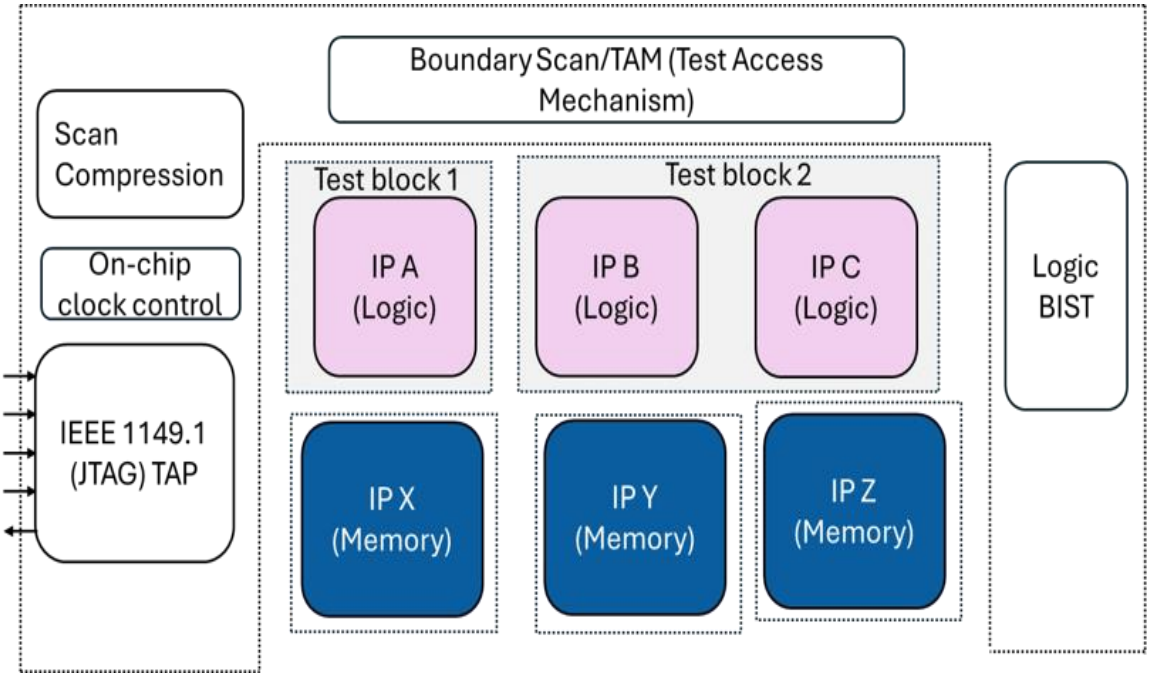

Figure 5: Static DFT: partitioned IPs, lock-up guarded, balanced scan chains

Block-partitioned SoC test architecture reflects a grouping of logic IP (A-C) and memory IP (X-Z) under a DFT regime with scan chain continuity, ordered stitching, and lock-up latches applied to asynchronous or skewed crossings, as shown in the figure above. Scanenable is single-source and glitch-free, orthogonal with functional enable, and scan clocks are gated in mission mode. The ports are connected to corresponding decompressor/compactor channels by a bijective map, and chain lengths are aligned to meet time budgets on the testers. A universal parameterized property library is size-scalable with respect to chain count, decompressor width, and mask topology.

In addition to linting, a formal verification shows that the DFT wrapper and the TPG could not oppose safety conditions. The assertions of scan-enable non-overlapping with functional capture, decompressor enables, and X-mask enables, satisfying mutual-exclusion requirements, and chain-heads and tails legal sinks and sources permanently connected in test mode. To address timing hazards, sequencing properties are extended with legal changes in functionality (to scan-shift and scan-capture and back) bounded by latency and described on control nets. Insertion of scan and TPG logic, then passes equivalence checking to ensure that functional behavior is not altered when test mode is detected [25].

4.2 Dynamic Simulation & Fault Simulation

Realistic shift and capture sequences and detectability are exercised in dynamic checks that validate detectability. Under realistic tester limits today, Gate-level simulation respects tester-equivalent clocking: separate scan and functional clocks, ratioed frequencies, duty-cycle restrictions, and explicitly defined hand-off between shift and capture. The resets and scan controls are pulled low with respect to the capture pulse to meet setup and hold requirements across corners. Each metastability window and glitch propagation on a mode multiplexer and in clock gates is revealed by the mode-multiplexer and clock-gate simulations in Standard Delay Format min/max to avoid spurious captures.

The explicit modelling of the X-propagation is included. Uninitialised flops, tri-stated buses, mixed-signal boundaries, and powered-down islands are considered unknown. The harness separates pessimistic and optimistic X modes and measures the difference between their respective levels of observability so that masking strategies may be justified. Illegal-state constraints guard against the triggering of damaging combinations, and scan-compression constraints restrict pattern encodings to values realisable on silicon. Power-aware test is imposed by observing pattern toggle density per pattern, maximum switching across sliding windows, and V/V activity; patterns that violate IR or thermal wraps are vetoed.

The simulation of faults determines the quantitative aspect of the detectability of the target models. THRBug-S combines several stuck-at and transition faults per pass using reconvergent fan-out compression in bit-parallel engines, and sets enforceable sensitisation rules with two-cycle launch/capture stimuli aligned to the intended at-speed scheme. Pattern compaction is trackable and hence detection sets are maximal without breaching clock, power, or decompressor constraints [36]. Workloads are split by block, clock domain, or chain group; caching of fault effects and early-stop on detection minimizes wall time.

4.3 Emulation / Prototype-in-the-Loop

Where sequence length, realistic unknown behaviour, or long-run power interactions are beyond the capabilities of software, the TPG, phase shifters, mask controllers, and MISR are implemented in an emulator or FPGA prototype. A simulation of a scan clock pretreats the decompressor and phase shifters; capture clocks to the device under test simulate phase correlations and clock gating timings. MISR taps are routed to debug bridges and rolling trace buffers diagnostics to capture selected scan segments before and following captures to diagnostic pattern efficacy and X-escape paths.

Vietnam launches synchronisers to accommodate rate mismatches between emulated shift and capture clocks to ensure launch/ capture phasing is appropriate to the automated test device intent [29]. At-speed validity is limited due to the lack of cycle-accurate timing in emulators, to control-sequencing verification and long-run-zd behaviour; at-speed path-delay-sensing assertion claims are made only in SDF-timed software or on-chip at-speed capture. Checks that are performed are signature-consistency checks against identical runs with identical seeds, fault-injection stubs, and test checks that ensure observability paths are truthfully transversed end-to-end.

4.4 Statistical Confidence & Aliasing Analysis

Sign-off needs to measure uncertainty and residual risk, rather than providing single-point estimates. The coverage confidence intervals calculated by the framework are based on non-parametric bootstrap resampling at the pattern level and, in the case of the logic built-in self-test, on the seed level as well. In all programmes that are repeated across designs, a hierarchical bootstrap retains within-design correlation before aggregation, resulting in intervals that represent deployment variance. Reports will include counts at fixed patterns of coverage and at equal tester-time budgets, so that reported gains will not be due to numbers at different budgets.

The phenomenon of aliasing is investigated in two directions: The theoretical upper limit is based on the number of inputs in signature registers n, with ideal alias probabilities of about 2^−n under random error syndromes; scaled to the customer parts-per-million requirements, limits per compaction stage are calculated. Empirical injection also includes systematically associated fault families, on-and-off profiles that correspond to systematically arranged defects, and deterministic uncertainties to estimate non-random syndromes. The observed collision rate is compared to the bond; in case there is an alternation of excess collisions, then the plan narrows the MISR width or phase decorrelates the phase, or changes the X-masking placement. Seed-to-seed variance and a stop rule ending reseeding when incremental coverage drops below a certain level with high confidence are reported.

4.5 Automation, CI/CD, and Reproducible Sign-off

Deterministic automation and constraint-aware scheduling are necessary in implementing the framework. The pipeline is represented as a directed acyclic graph that takes the static checks, formal proof, simulation, fault simulation, emulation, and statistical reporting. All inputs, libraries, netlists, SDFs, constraints, fault lists, decompressors, MISR parameters, seeds, versions of the tools, and command options are cryptographically hashed into a manifest that is embedded in all reports [8]. Container images, or the use of virtual environments, lock up EDA dependencies via version pinning or seed naming to ensure bit-identical reruns.

DFT sign-off, as shown in the figure below, is a directed acyclic graph orchestrated by a deterministic CI/CD pipeline featuring static checks, formal proofs, simulation, fault simulation, emulation, and statistical reporting. All inputs--libraries, netlists, SDFs, constraints, fault lists, decompressor and MISR parameters, seeds, tool versions, and command options--are cryptographically hashed into a manifest submitted in reports. Version-pinned runners and containerized environments ensure the reproducibility of the environment and version of the EDA, making it possible to run identical bits through build and release, automated packaging, testing, and publication, with guard gates across the rerunnable project lifecycle.

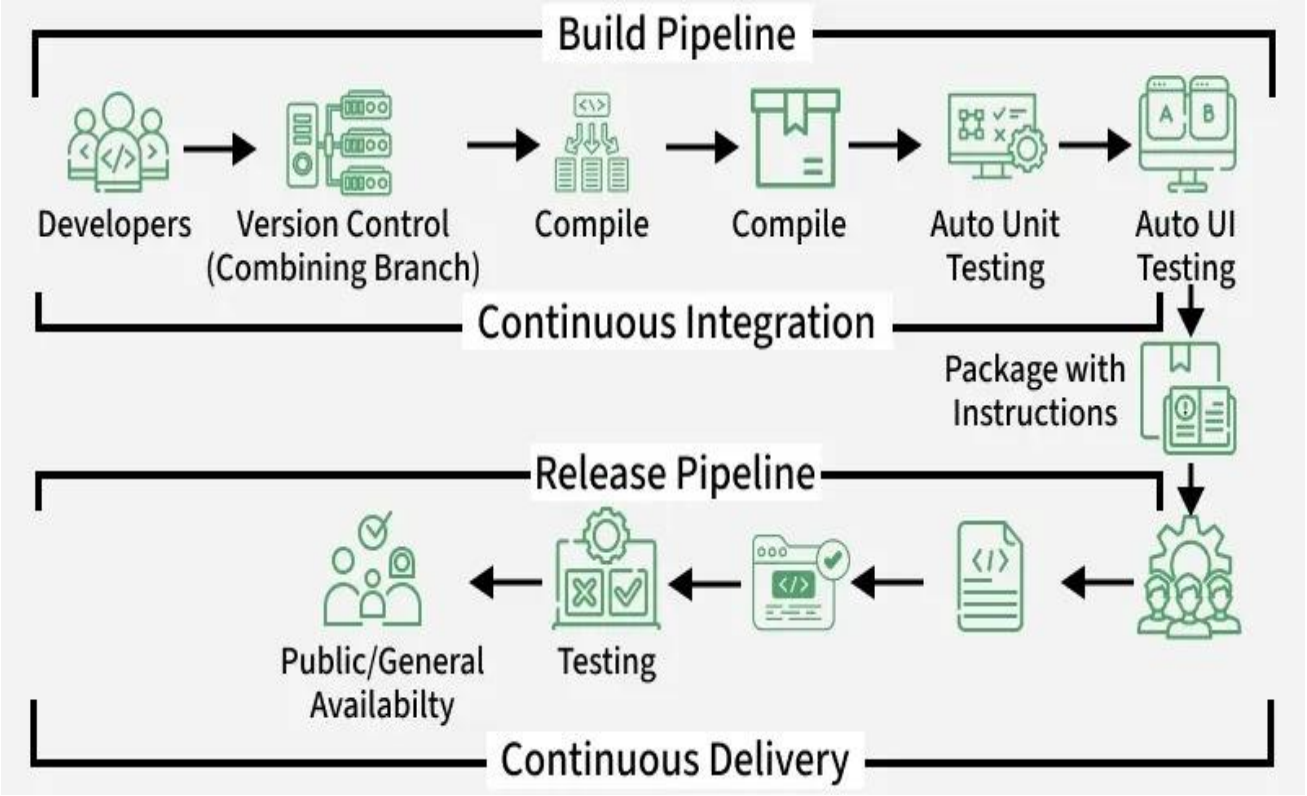

Figure 6: Deterministic CI/CD pipeline for reproducible DFT sign-off artifacts

Job orchestration focuses on optimizing the number of uncertainty reduction per compute-hour and observes the various resource constraints like shortages of fault-simulation licenses and emulators, as in the throughput optimizations on other constrained operations [22]. Sign-off data in machine-readable reports lists metrics, confidence intervals, alias analysis, dangerous pattern veto logs, the manifest, supporting audits, and real-time compliance. Dashboards present data in the form of numeric tables and graphical snapshots of waves and signature heatmaps; compared with graphical displays as well as text-based summaries, they allow humans to be more efficient in their analysis effort [33].

5. Experiments and Results

5.1 Experimental Setup

The experimental campaign compared three manufacturing scaling decades of 28 nm, 16 nm, and 7 nm with a range of test-power and scan architectures to span a breadth of technologies. Design D1 (MCU-class subsystem) has 5.1 M logic gates, 0.35 M scan FF, 320 scan chains, and lock-up latches in two asynchronous clock domains, as highlighted in Table 1 below. D2 (16 nm networking accelerator) has 18.9 million gates, 1.2 million scan flip-flops, 960 chains per domain, each with level shifters and an isolation cell. D3 (a 7-nm application processor cluster) has 52.4 million gates, 3.8 million scan flip-flops, and 2,400 chains in a domain using retention and power-gate islands. It used a commercial ATPG tool to generate deterministic patterns that use dynamic compaction and X-fill policies (biased, adjacent, power-aware). A fault simulation was used on stuck-at as well as transition and sampled path-delay lists; the same HDL was again X-propagated using X-propagation with pessimistic semantics and user-supplied black boxes to cover analog macros and SRAM peripheries.

Table 1: Experimental setup summary for D1–D3 designs (28–7 nm)

Parameter	D1 — 28 nm MCU-class	D2 — 16 nm networking accelerator	D3 — 7 nm application processor

Node / class / scale	5.1 M gates; 0.35 M scan FFs	18.9 M gates; 1.2 M scan FFs	52.4 M gates; 3.8 M scan FFs
Scan chains (per domain)	320	960	2,400
Domains & DFT features	2 async domains; lock-up latches on crossings	Multi-domain; level shifters + isolation cells	Multi-domain; retention flops; power-gate/voltage islands
ATPG	Deterministic; dynamic compaction; X-fill: biased/adjacent/power-aware	Deterministic; dynamic compaction; X-fill: biased/adjacent/power-aware	Deterministic; dynamic compaction; X-fill: biased/adjacent/power-aware
Fault sim & X-prop	Parallel engine; SA/transition/sampled path-delay; pessimistic X-prop; analog/SRAM black-boxes	Parallel engine; SA/transition/sampled path-delay; pessimistic X-prop; analog/SRAM black-boxes	Parallel engine; SA/transition/sampled path-delay; pessimistic X-prop; analog/SRAM black-boxes
Timing & power sign-off	At-speed offset −25 MHz; sliding-window toggle + IR models (calibrated)	At-speed offset −50 MHz; sliding-window toggle + IR models (calibrated)	At-speed offset −50 MHz; sliding-window toggle + IR models (calibrated)
LBIST & MISR	LFSR 32/48-bit; phase shifters 8×; reseed every 2K; MISR {32, 48, 64}	LFSR 32/48-bit; phase shifters 16×; reseed every 2K; MISR {32, 48, 64}	LFSR 32/48-bit; phase shifters 24×; reseed every 2K; MISR {32, 48, 64}
Compression / tester / budget & gates	Decompressor 64/96/128 ch; expansion 8×/12×/16×; virtual tester 512 ch, 64 M vectors, 2 GB/s; 1,800 s; gates: coverage, alias vs n, toggle/IR compliance	Decompressor 64/96/128 ch; expansion 8×/12×/16×; virtual tester 512 ch, 64 M vectors, 2 GB/s; 2,400 s; gates: coverage, alias vs n, toggle/IR compliance	Decompressor 64/96/128 ch; expansion 8×/12×/16×; virtual tester 512 ch, 64 M vectors, 2 GB/s; 3,600 s; gates: coverage, alias vs n, toggle/IR compliance

At-speed capture timing mode applied a timing scheme with -25 MHz (D1) and -50 MHz (D2, D3) bounds of scan and shift. Power sign-off at test was based on per-pattern vector activity: sliding-window toggle budgets and IR-drop models calibrated at sign-off parasitics. LBIST used 32 and 48-bit LFSRs with primitive polynomials along with per design phase shifters (8x in case of D1, 16x in case of D2, and 24x in case of D3) and reseeding after every 2K patterns. In MISRs of $\in$ {32, 48, 64} signatures were recorded. Compression tests were applied using 64, 96, and 128 channel decompressors with 8x, 12x, and 16x internal expansion. The virtual tester simulated 512 data channels, 64 M vector depth, 2 GB/s upload bandwidth, and capture-cadence constraints. Gate coverage targets, aliasing correction in accordance with n, compliance with toggle/IR constraints, and overall test time budgets of 1,800 s (D1), 2,400 s (D2), and 3,600 s (D3), were the acceptance gates.

5.2 Baselines

There were three baselines for comparison. The unconstrained deterministic ATPG set a strict upper bound on the raw controllability/observability; patterns were compacted perfectly, but it allowed illegal functional states as well as unconstrained switching [2]. Such vectors cannot be played on physical testers, but this

configuration is the baseline that captures the delta that realistic constraints inject. Uniform LBIST with a fixed 32-bit seed and no reseeding and unweighted fault placement was the configuration that allowed us to quantify the innate pseudo-random performances and, in fact, to form a chain on hard-to-control faults. A trim compression flow employed a 64-channel decompressor with non-adaptive expansion and fixed X-masking of known unknown-prone endpoints; the point was to isolate channel count and compaction benefits with no advanced masking or power-conscious fill. To remain comparable, equal clocking assumptions, legal-state restrictions, and tester pin maps were used where possible. The results characterizing outcome were: (i) stuck-at, transition, and sampled path-delay fault coverages; (ii) the total pattern count and the tester time of channelization, expansion, and capture cadence; (iii) toggle and IR-drop power compliance checks; and (iv)empirical an estimate of the aliasing provided by the MISR width n with independent seeds collision search. The baselines were limited to the same time budgets as the principal flow, and vectors beyond toggle limits were marked unsafe (instead of being dropped silently) to count the unsafe-vector rate. LBIST baselines extended to 1.2 million patterns or where the marginal detection rate was less than 1e-6 per pattern.

5.3 Main Results

The coverage-volume curves portrayed diminishing returns, but the Pareto front was enhanced when realistic constraints, compression, and reseeding were used in combination. On D1, the unconstrained deterministic baseline was able to solve 99.2% stuck-at and 94.8% transition at 128k patterns, but 23% of vectors exceeded the toggle envelope. The imposition of legal-state, CDC, and power restrictions led to a decrease in transition coverage, with 93.9 percent still being covered and unsafe vectors removed [4]. Compared with adding 96-channel decompression that covered 99.2% and 93.8%, and resulted in a 38% decrease in the pattern volume, the total test time fitted within the 1800 s budget, having a 0.4% margin of the vectors. Using n = 32 MISR, uniform LBIST recorded 97.1 percent stuck-at and 89.6 percent transition after 1.2M patterns on D2. The addition of reseeding within 2k patterns and an increase in n = 48 achieved a transition to 91.3% with a test-time improvement of 25% when compaction and the capture cadence were incorporated.

Table 2: A summary of main results—coverage, pattern counts, toggle compliance, and MISR selections for D1–D3

Scenario	Stuck-at (%)	Transition (%)	Patterns	Compliance / Notes
D1: Deterministic ATPG (unconstrained baseline)	99.2	94.8	128k	23% of vectors exceeded toggle envelope
D1: Deterministic ATPG (with legal/CDC/power constraints)	99.2	93.9	N/R	Unsafe vectors removed; coverage drop due to constraints
D1: 96-channel compression	99.2	93.8	≈79k	−38% patterns vs. baseline; test time ≤ 1,800 s (≈0.4% margin)

D2: LBIST, uniform, MISR n=32	97.1	89.6	1.2M	Baseline LBIST before reseeding/compaction
D2: LBIST + reseeding, MISR n=48	N/R	91.3	N/R	Test-time improved by 25% with compaction & capture cadence
D2: Deterministic ATPG + 96-channel compression	98.9	93.5	360k	99.2% vectors within toggle limits
D3: Deterministic ATPG, 128-channel, 12× expansion	98.7	93.1	480k	Meets coverage with fewer patterns than 64-channel
D3: Deterministic ATPG, 64-channel	98.7	93.1	690k	More patterns required to match 128-channel coverage
D3: LBIST, 24× phase shifters, reseeding	N/R	90.4	1.6M	Pre-weighted pseudo-random baseline
D3: LBIST + weighted pseudo-random	N/R	91.0	1.3M	Higher transition coverage with fewer patterns
Power-aware fill effect (all designs)	N/R	N/R	N/R	Over-budget vectors reduced from 8–23% (naive) to ≤1% with power-aware fill + cycle stretching
X-masking study (D2)	−0.2 pp	−2.1 pp	N/R	Mask off: up to −14% usable capture bits; selective bitwise X-blocking recovers observability

Aliasing behavior & choices	N/R	N/R	N/R	n=32 ⇒ ~2.3e-10/signature; 0/10M collisions; chosen MISR: D1=32, D2=48, D3=48

Deterministic ATPG with 96-channel compression had 98.9 percent stuck-at and 93.5 percent transition coverage, 360k patterns, with 99.2 percent of the vectors remaining below the toggle limits. On D3, 128 channels of deterministic ATPG using 12x internal expansion achieved 98.7% coverage of faults and 93.1% coverage of transitions at 480k patterns; 64 channels of ATPG required 690k patterns to reach the same coverage levels. Phase shifting at 24x and reseeding reached 90.4 percent transition at 1.6 million patterns, and the weighted pseudo-random distribution further boosted the transition to 91.0 percent at 1.3 million patterns. Power-aware fill provided a practical increase in toggle compliance. Vectors over the sliding-window budget dropped 8-23 percent under naive fill, and to 1 percent or less with power-aware fill and smart cycle stretching. The effects of X-masking on the observability were singled out on D2 through paired runs.

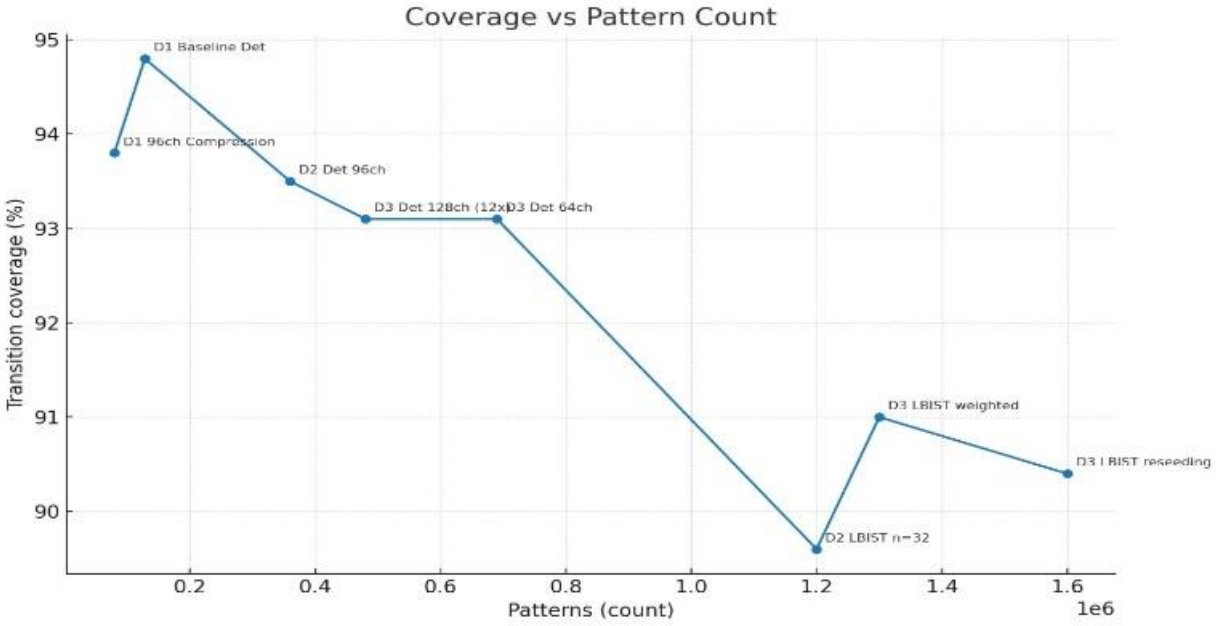

Figure 7: A line-graph showing transition coverage vs. pattern count across D1–D3 scenarios

When masking is turned off, the convergence of unknowns in analog wrappers, as well as uninitialized peripheries of SRAM, reduces the usable capture bits by up to 14%, with the 2.1 percentage point effect on the transition coverage. Selective bitwise X-blocking, which allowed observability to be regained with only a 0.2 percentage point loss in stuck-at coverage, proved better than coarse gating. Aliasing behavior tracked theory: for n = 32, $1/2^n \approx 2.3e\text{-}10$ per signature; no collisions were observed in 10 million random-seeded trials. For n = 48 and n = 64, empirical rates were below tester observability. Consequently, n = 48 was selected for LBIST on D2 and D3, while n = 32 sufficed for D1.

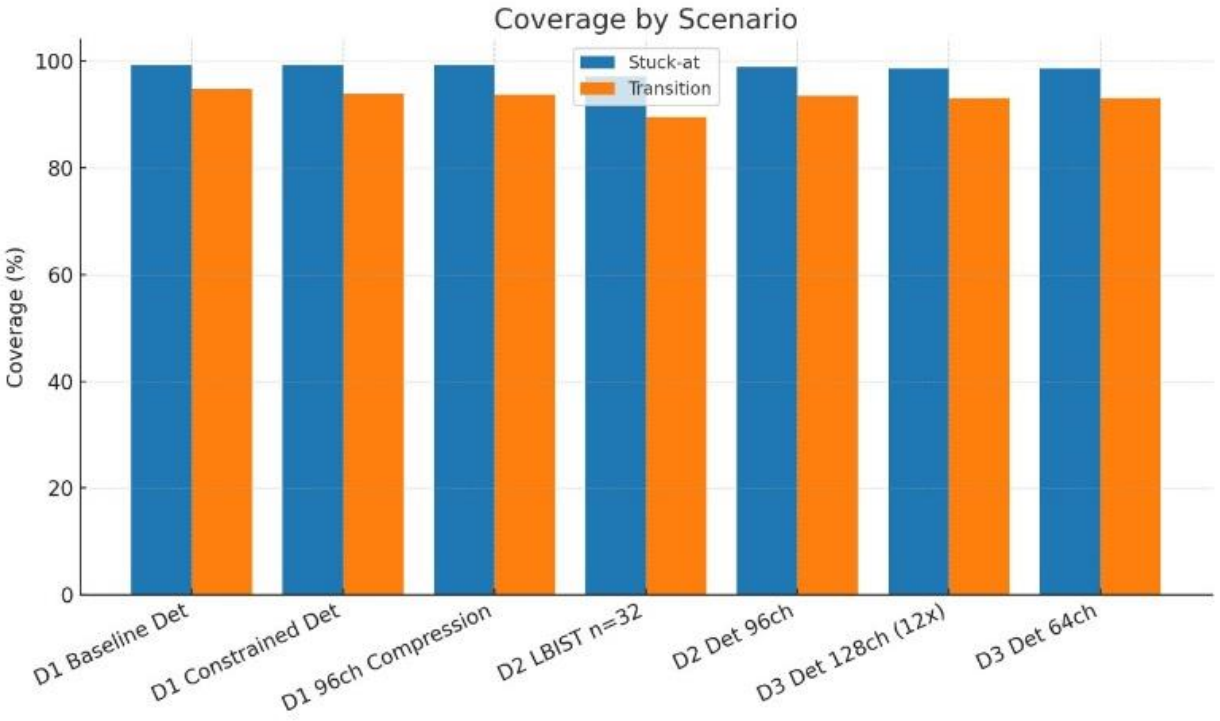

Figure 8: A graph showing stuck-at and transition coverage by scenario (D1–D3).

5.4 Ablations

Fidelity had a first-order impact on both depictions of coverage and deployability. On D2, when six unsafe debug modes were deliberately left out of the constraint set, apparent transition coverage increased by 0.7 percentage points, but 4.6 percent of the patterns were discarded because their legal stimulus traces were marked as illegal by the tester. The reinstatement of constraints brought the reported coverage down to the realistic figure, and it provided full playability for testers. As shown in the Table 1 and Figure 9, the propagation mode also had an effect: optimistic X propagation yielded 1.3 to 1.9 points higher transition coverage than pessimistic modeling across designs in sync with black-boxed macros. The pattern volume and observability were affected by the decompressor topology.

Table 3: A summary of ablation factors, quantitative effects, and practical trade-offs

Ablation factor	**Change / setting**	**Quantitative effect**	**Practical note**
Constraints fidelity (D2)	Omit 6 unsafe debug modes vs. reinstated	Apparent transition **+0.7 pp**; **4.6%** vectors flagged illegal; reinstatement returns to realistic coverage, **0%** illegal	Keep full legal-state/CDC/power constraints for tester-playable vectors
X-propagation mode	Optimistic vs. pessimistic	Optimistic reports **+1.3 to +1.9 pp** higher transition coverage	Use pessimistic with macro black-boxes for sign-off realism
Decompressor topology (D3)	64 to 96 channels, 12× expansion	Pattern count **−22%** at fixed coverage	Good trade-off; stays within ATE I/O
Decompressor topology (D3)	96 to 128 channels	Additional **−9%** patterns; ATE I/O exhausted; time-division with memory test	Lowers patterns but increases overall test time

LBIST phase-shifter depth (D2)	Baseline to **+8×**	Coverage **+1.1 pp** at same pattern count	Modest area/routing increase
LBIST phase-shifter depth (D2)	**+8× to +16×**	**+2.0 pp** additional (≈ **+3.1 pp** vs baseline); area **≈3×**; routing complexity increases	Only justify when coverage gap demands it
Seed cadence	Reseed every **2k** vs. **≥8k** patterns	**≥8k** leaves **3–5%** more residual transition faults after **1M** patterns	Prefer **2k** cadence for early detection
Weighted pseudo-random + fill	Enable weighting with power-aware fill	Controllability increases; toggling increases slightly; compliance **preserved**	Use weighting + power-aware fill to balance coverage and power

Migrating to 64-96 channels with 12x expansion on D3 cut pattern count by 22% at fixed coverage; up to 128 channels cut a further 9 percent but exhaust the test set I/O, necessitating time-division tactics with memory test, which increases overall test time. Phase-shifter depth in LBIST balanced pattern independence with area: an 8x increase gained 1.1 points in D2 pattern coverage at the exact pattern count, and a 16x increase, two more points at triple the area, and an increase in routing complexity in the DFT shell. Pseudo-random efficiency was influenced by seed cadence—factors reducing early detection. Seeding every 2k patterns maximized early fault detection; longer cadences (≥8k) reduced pattern diversity and left 3-5% more residual transition faults after 1 million patterns. Weighting made controllability in low-observability cones easier to control, but increased toggling a little; however, weighted power-aware fill preserved compliance.

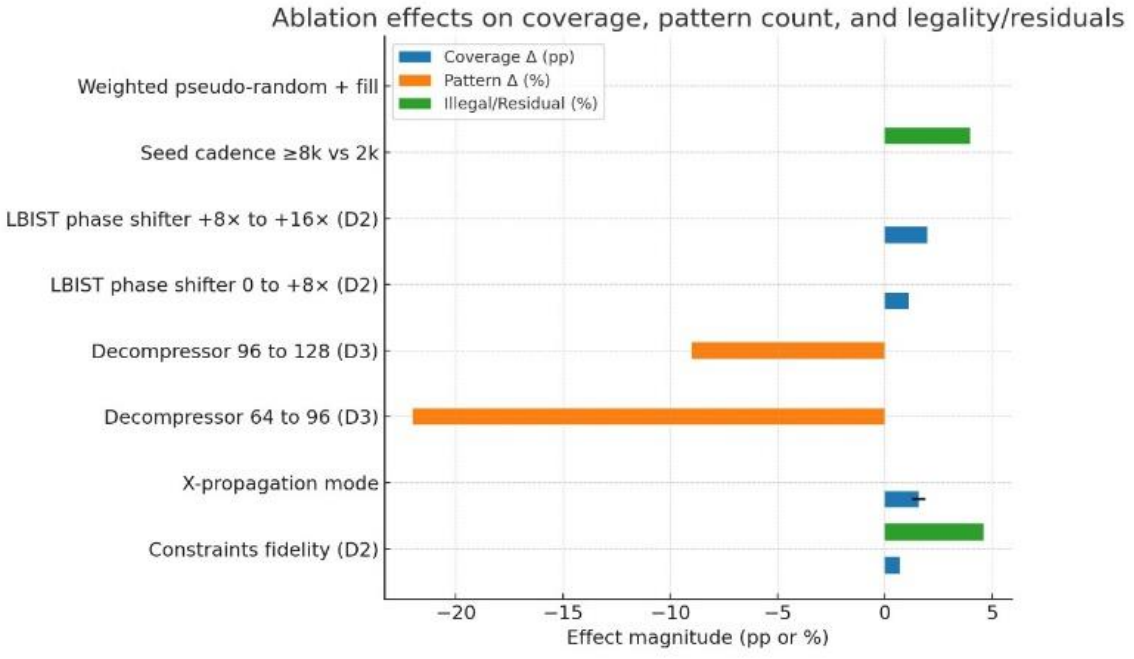

Figure 9: Ablation impacts on coverage, pattern count, and legality/residual metrics

5.5 Robustness & Error Analysis

Resistance to re-randomization of seeds and sequencing of patterns, constraint noise, and increase of X-density

were also evaluated. Across configurations, experts used five independent trials with different random seeds and pattern scrambling of each of the trials. On D2, the standard deviation of stuck-at coverage using reseeded LBIST was ±0.18 points and transition coverage ±0.22, ordering effects were insignificant following dynamic compaction since late patterns cover residual faults with low structural overlap. The noise of constraint was more poisonous. Adding 5 percent random flips to the constraint set overstated transition coverage by as much as 1.1 points and tripled tester rejects; an audit that cross-checked encoding of unreachable states removed the effect. Using Rising X-density, it was possible to activate more analogy macro stubs and postpone SRAM initialization. A 10-percent increase in X-density decreased the transition coverage by 0.9 to 1.4 points, unless selective per-capture X-blocking was used. A two-level mitigation was found to be effective: increase mask granularity at potentially realistic unknown sources, and also impose a hard limit on masked bits per capture to guard against stuck-at visibility. A lightweight dynamic-memory priority index was used in conjunction with signature clustering to maintain context in long fault-simulation logs and to identify recurring motifs of failure [28]. The key residual failure modes were correlated to reconvergent fan-out with deep sequential depth and to cross-domain scan stitch near isolation cells; localized fill directives and clock-domain-aware compaction addressed both.

6.1 Implications for Industrial Flows

Integrating the TPG verification into the silicon development lifecycle requires precise placement, clear ownership responsibilities, and an exit criterion. At RTL, enforceable constraints, scan-enable exclusivity, legal test-mode multiplexing, and X-robust coding conventions are represented as assertions and lint rules, causing the build to fail when these violations occur. The structural checks of chain-scan continuity, lock-up latches to handle inter-clock crossings, and controllability/observability guards on inserting DFT are checked automatically; equivalence checking ensures that scan structures and TPGs do not change functional intent.

Timing sign-off should follow place-and-route and must represent test-mode corners, scan-shift and capture timing exceptions, as well as SI checks under worst-case switching. Sign-off assembles a machine-readable portfolio: a confidence interval coverage, an alias-risk assessment, power envelopes per domain, constraint verification, and a manifest compliant with tester memory. Handoffs to test engineering start at synthesis freeze, allowing early pilot loading of the ATE and channel allocation and vector memory fit feedback. The ability to perform RTL validation via wrapperized access and modular pattern reuse eases integration of SoCs and optimizes debug time; verification of these checks early under top-level assembly makes ownership clear and unchurnable [7].

6. Discussion

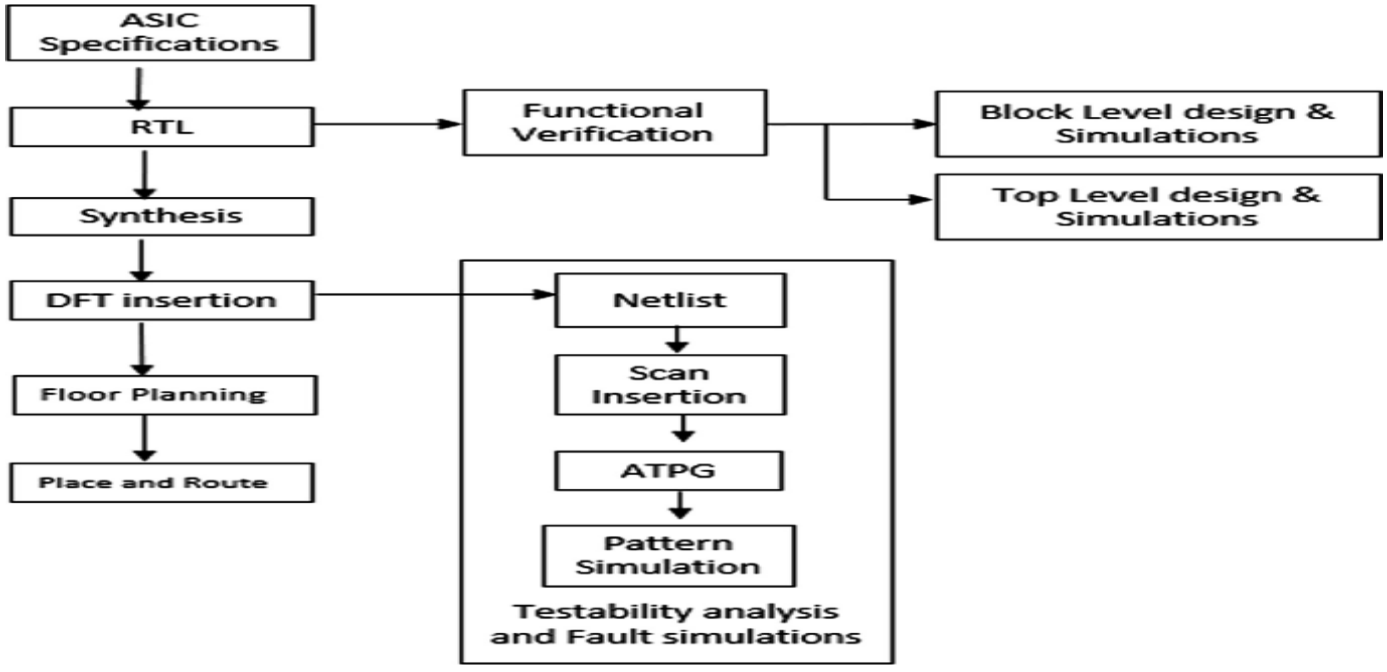

Figure 10: End-to-end ASIC flow with integrated DFT/TPG verification and sign-off

As shown in Figure 10 above, TPG verification is inserted throughout the silicon flow, including RTL assertions and lint, to enforce scan enable isolation and legal multiplex test-mode [12]. DFT insertion generates scan chains and lock-up latches, and the equivalence flow maintains functional intent—timing sign-off. Then there is shifting/capture corner analysis and worst-case SI. A machine-readable portfolio-coverage confidence, alias risk, power envelopes, and constraints help enable early ATE pilots at synthesis freeze to wrap and validate RTL in a modular fashion and reuse patterns at scale.

6.2 Key Trade-offs

Verification needs to bring to the surface the true trade-space rather than relying on a single metric. Spreading coverage versus pattern count is the first dimension: dynamic compaction and wider decompressor channels reduce data volumes, but aggressive compaction may increase switching bursts and can cause IR violations. Applies a practical approach using per-pattern toggle histograms, imposes sliding-window ceilings on mean and peak activity, and blocks outliers. The second axis is MISR width versus area: the aliasing probability of linear signature analysis scales inversely as 2^−k, a k-bit signature, but short-run transients and non-uniform error models tend to deviate at small k; engineers size MISRs by targeting a ppm performance, and then testing the off-design performance with stress tests including correlated response classes and with sequence length [9, 35].

A third axis is decompressor channels against I/O limits. Additional channels enhance controllability and minimize the pattern volume, but use pins and routing resources, and may inhibit timing. EDT-style architectures are not only economic at high compression ratios, and they retain determinism, but they are also subject to rigorous constraint handling to prevent illegal states, careful X-masking or X-bounding to avoid the loss of observations, and corner cases of the mask controllers have to be executed explicitly. Policies ought to be biased towards legality and power envelopes, rather than pure compaction: reject patterns that push toggle budget or thermal limits even when they give a slight improvement to coverage. When memory is scarce, compaction and channel count are optimized along with scan-chain balancing to avoid compromising detection quality with too many bursts [27].

6.3 Toolchain Integration Patterns

Long-term quality relies on automation and reproducibility. Teams have versioned libraries of property-encrypted property, scan-enable one-hotness, exclusive modes of test, decompressor safety, mask legality--and maintain DFT lint rulepacks with taxonomies of waivers. Continuous-integration gates consist of: (1) structural checks right after DFT insertion, (2) formal proofs at block and cut levels with bounded-time obligations, (3) nightly X-aware gate simulations with compressed patterns with per-pattern power screens, and (4) weekly fault-simulation snapshots to track coverage and alias-risk.

JSON, CSV Artifacts are exported to dashboards: per-MISR alias-ppm-estimates, per-domain decompression density, a sliding-window toggle, and per-scan-cell-plus-per-decompression-slice-keyed triage table. Reproducibility is dependent on tool and seed pinning, artifact hashing, and reproducible builds; any tool change corresponds to an A/B run use. Low-power test plans are here part and parcel of policy: pattern throttling, chain partitioning, and capture shapers are chosen automatically to achieve IR-drop constraints without losing detection targets.

6.4 Generalizability across Designs & Nodes

The architecture should be device transferable between IP vendors, technologies, nodes, and power architectures. The IP boundary also provides standardized access networks to embedded instruments, and this decoupling of local implementation to the SoC integration enables repeatable test activation, status capture, and debug without custom glue logic. Use of an instrument-access standard enables verification collateral, such as properties, checkers, scripts, and test templates, to transfer among products without modification and to work side-by-side with boundary-scan. Such portability is ideally applicable when chip partitions involve third-party cores, safety islands, and always-on domains; when these partitions are combined, a common access description eliminates bespoke wrappers and reduces re-verification tasks [24].

Sensitivity to corners of process, voltage, and temperature increases at advanced nodes; thus, timing closure in test-mode must also enumerate corners and derate to the ATE timing set. Statistical timing monitors should be used to identify marginal hold or setup in a long scan chain,

especially across voltage-translated boundaries incorporated into test-mode simulation. To ensure generalization, many reports distinguish design-intrinsic effects (e.g., scan depth distribution) and environmental effects (such as temperature dependence of capture windows), and document calibrations such as slower scan clocks in cold corners. Document scan-capture modes are now explicitly defined.

6.5 Future Considerations

Four practical steps as follows deserve future verification attention: Stop rules about the convergence of coverages and confidence of alias must regulate adaptive seed and reseed policies of LBIST and weighted pseudo-random modes [19]. Every seed keeps a running detection curve; when the slope drops below a set threshold and when the confidence interval on incremental coverage intersects with zero, the seed cancels out. MISR alias risk is monitored as a moving ppm estimate; seeds can only be escalated when the coverage gain per millisecond of test time exceeds a budgeted minimum and the alias bound tightens. The Default X-propagation fidelity would be shifted to contract-based verification. Pessimism is mitigated with selective X-bounding at known origins of Xs, and bejewelled with formal properties that establish the existence of mask assertions only under traceable preconditions. Optimism is limited by deterministic initialization, CDC-aware reset sequencing, and assertions that prohibit illegal don't-care interpretations in scan and capture modes. Tooling is expected to produce X-origin traces, mask-condition coverages, and waiver differences per build to make decisions auditable.

Efficient scheduling and power-aware scheduling move to the top-tier automation targets. Pattern ordering also seeks to minimize droop by alternating high-activity and low-activity slices, phasing the scan-clock, and the placement of capture-strobes with per-domain IR limits and tester timing. The scheduler uses machine-checkable contracts, maximum burst length, capture density, sliding-window switches, cool-off spacing, and concurrency limits. It vetoes patterns that would increase coverage beyond the envelope, even as coverage increases. Per-pattern activity, deduced droop risk, and ATE memory fit are reported to facilitate rapid triage. Structured evidence that is packaged to meet automotive functional-safety requirements should be used to host safety-critical programs [14]. The evidence pack contains a traceability matrix of hazards to test goals, assumptions of use, formal property outcomes, coverage, and alias ppm with confidence bounds, power-envelope compliance, as well as reproducibility metadata (tool versions, seeds, hashes). The material enables review, facilitates change-impact analysis, and reuse across derivatives without re-inventing the verification workflow.

7. Limitations & Threats to Validity

This section lists significant methodological risks and practical safeguards in the context of TPG verification.

7.1 Internal Validity

Label noise is related to timing approximation or X-propagation simplification in gate-level fault simulation. Negative pessimistic X handling underestimates observability, and positive optimistic ones mask the reality of the underlying label, producing false positives and false negatives. Another internal risk is constraint mis-specification, as illegal-state filters, scan-mode clocking, power caps, and tester channel limits may be specified incorrectly and result in a pattern that is, in fact, infeasible but passes.

Testbench fidelity is also essential as the order of pattern application, scan-enable sequence, and reset release must match the tester sliding protocol. Mitigations used are X-prop settings tuned with small hand-proven blocks, dual-mode fault simulation (optimistic and pessimistic) with tagging of disagreements, constraint unit tests with golden pattern replay, and assertions to guard scan-mode exclusive completeness, mask enable legality, and decompressor safety. Expectations between properties and reports are reduced through traceability, which minimizes silent drift in the specifications as general advice on safety-critical systems [32].

7.2 External Validity

Results are not necessarily transferable across designs, nodes, or vendor flows. Diversity affects controllability and observability, and contributes to diversity. It can be low-power islands, retention flops, and CDC schemes and analog wrappers, which can increase X density and complicate the effectiveness of masking—the differences between place-and-route influence reconvergence, slew, and glitch windows, shifting transition/path-delay detectability [30]. Toolchains differ in scan stitching, decompressor topology, and X-bounding implementations, whereas safety classifications (such as ASIL levels) dictate different fault-coverage floors and diagnostic metrics.

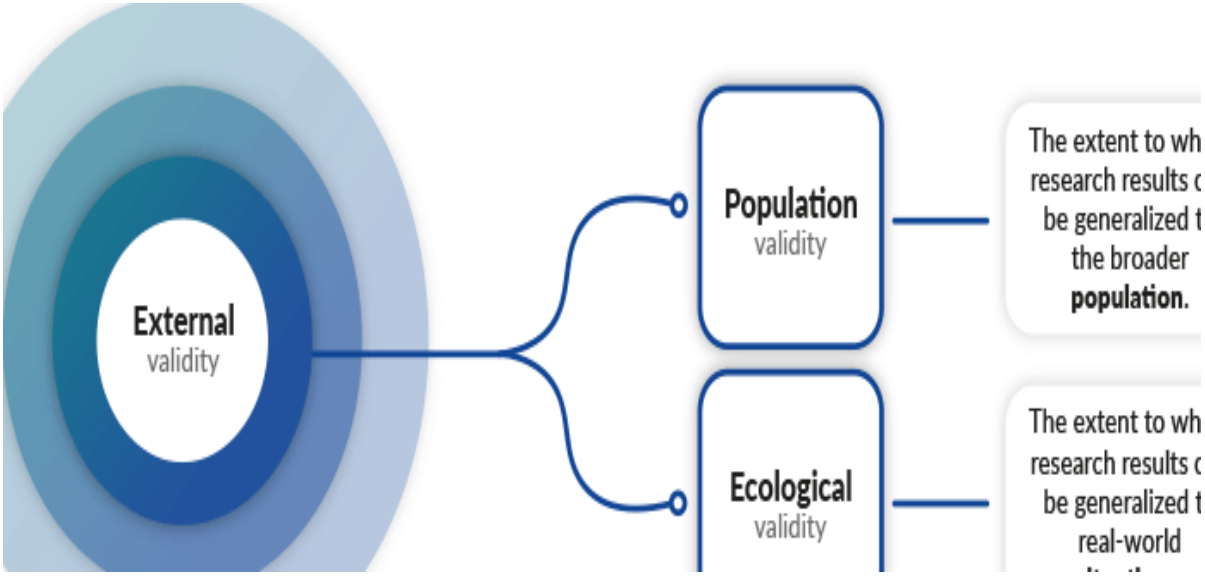

Figure 11: External validity types: population and ecological generalization limits

External validity has two sides of applicability toward TPG verification, as shown in Figure 11 above. The question posed by population validity regards whether the coverage and alias measures can be generalized to the broader population design, nodes, and vendor toolchains. Ecological validity raises the question of whether findings apply to actual tester configurations of low-power islands, retention flops, CDC schemes, analog wrappers, and variances in P&R that can change reconvergence, slew, and glitch windows. Tester variation adds additional drift: pin maps, available scan frequency, vector formats (STIL/WGL/EVCD), cycle accuracy, and per-pattern current limits all determine what may be run at speed. To check transferability, the verification flow must involve a design-holdout scheme (train/tune on a subset of designs, validate on a disjoint design family), design-specific power budgets, and tester-in-the-loop dry-runs that compile, time, and double-check patterns against the actual ATE system prior to silicon.

7.3 Construct Validity

The proxy gap between simulated metrics and the shipped quality is a prime example of a suppressive statute. The issue with high simulated stuck-at/transition/path-delay coverage is that it does not guarantee a low DPM when defect mechanisms are off, or you have defect mechanisms that deviate (resistive bridges, small-delay intermittency, cell-internal defects), or X-masking that hides marginal behavior. Aliasing analysis is also sensitive, as textbook MISR estimates ($\approx 1/2^n$) assume independence; correlated responses on phase shifters, unknown sources, compressor reuse, or poorly seeded LBIST can significantly increase the empirical probability of aliasing.

Realistic protection involves such cell-sensitive fault models of critical libraries; such defect-oriented sampling with promising nets; such alias stress tests that introduce families of responses with correlations to measure deviation of responses against theory. Besides, coverage should be complemented with the confidence intervals, the estimate of the escape rate, and power-conscientious safety tests. Post-silicon correlation would be needed to relate the predicted detection hot spots to actual ATE fail signatures, and update constraints and pattern policies when the proxy gap was observed [17].

7.4 Reproducibility

Causes of threats to deterministic regeneration include random seeds, tool/version drift, and environment differences. LBIST initialization, TPG reseeding, decompressor fill, and X resolution are all seed-sensitive; runs must be PRNG seed-pinned and seed streams-persistent, and seed to pattern mappings must be logged. Variations in EDA packages, library/PDKs, and synthesis flags alter logic structure, including scan stitching, and make baselines invalid. To achieve reproducibility, it is necessary to have fully scripted flows, containerized tool chains, and hash the netlist, constraints, properties, and pattern dumps.

Telemetry improves auditability as stage-by-stage metrics, run manifests, and immutable storage of vectors written by the tester can provide end-to-end traceability.

Operationally, such controls reflect principled telemetry and property-tracking systems, where identifiers are consistent and communication channels can be trusted to deliver repeatable results [23]. A Release note should lock the toolchain, PRNG seeds; a regeneration is only accepted when results produce identical scan-in/scan-out waveforms.

8. Conclusion

This work provided an insightful illustration of how the task of test pattern generator (TPG) verification of modern and diverse system-on-chips is multi-objective in nature and represents a multi-objective optimization problem that involves the combination of controllability/observability, aliasing risk, X-propagation fidelity, tester constraint, and power integrity. This consideration of verification as an exercise of creating traceable evidence rather than a vaguely coupled set of checks, the work is co-aligned with deterministic ATPG, LBIST, and compressed-scan generators together with stated objectives (coverage, pattern volume, tester seconds, alias parts-per-million, and safe toggle/IR envelopes) validated within the confines of typical clocking, reset, and mode constraints. The outcome of this is a repeatable, auditable process that transforms verification of DFT into a defensible sign-off process.

Methodologically, it was able to insert structural DFT lint and parameterized property libraries along with equivalence checking to maintain functional behavior after insertion. The simulation levels include gate-level with paired optimistic/pessimistic X-prop modes that bounded observability uncertainty and generated selective X-masking or X-bounding and fault simulation quantified stuck-at, transition, and sampled path-delay coverage using legal stimulus and tester timing. Emulation was prototyped with the ability to emulate and capture the seed-to-seed variance. Statistical sign-off employed bootstrap confidence intervals (patterns and seeds) and two-level alias analysis: a mathematical 2-n limit on MISR width selection and a practical collision pressure on correlated response families. CI/CD automation, seed/version pinning, artifact hashing, and machine-readable manifests were used to provide reproducibility and auditability across toolchains.

The three-node experimentation can determine practical gains/ trade-offs. On D1, stacking compression to 96 channels maintained 99.2 and 93.8 percent stuck-at and transition coverage, reducing the pattern volume by 38 percent and fitting the budget of 1,800 seconds to the tester. Unconstrained vectors that exceeded toggle limits (approximately 23%) were removed by power-aware fill and routing. D2, LBIST reseeding alongside 48-bit MISR enhanced the transition coverage compared to the uniform LBIST test, and with compaction and cadence control, it reduced test time. On D3, the reduction in decompressor channels (64 to 128) at a fixed coverage compressed both environments, but tightened the I/O budget. Selective, bitwise X-masking restored observability at a cost of less than 0.2 percentage points of negative stuck-at coverage and outperformed coarse gating. Empirical alias rates were consistent with theory for $n \in \{32, 48, 64\}$; n=48 provided a robust margin for higher-complexity designs.

Based on these results, a more realistic sign-off contract is implied covering (1) minimum coverage per fault model with confidence intervals; (2) alias ppm thresholds mapped to MISR width with correlated-family stress results; (3) per-pattern and sliding-window toggle/IR checks with veto and rescheduling; (4) verified constraint legality (scan-enable exclusivity, decompressor/mask safety, clock-mode sequencing) and functional equivalence; (5) ATE fit (channels, vector memory, frequency) verified in in-silicon simulation pass; and Such artifacts facilitate clean handoffs to the test engineering standards and also ease bring-up, comparing the pre-silicon predictive information with ATE reality.

The study does not ignore the remaining ongoing risks and mentions possible future steps such as achieving X-prop fidelity at sources of known unknowns, applying adaptive LBIST stop-rules based on incremental coverage confidence, advancing energy-saving schedulers that interleaved high/low-activity bursts within domain envelopes, and bottling safety-case evidence (coverage metadata, alias, power, and reproducibility metadata) into ASIL programs. These practices allow TPG verification to be disciplined and metrics-driven, scaling with increasing design complexity and reducing test cost, and improving outgoing quality without sacrificing safety or silicon viability.

References

[1] Angadi, A., Umarani, S., Sajjanar, T., Karnam, R., Marali, K., & Shirakol, S. (2023, June). Architectural Design of Built in Self-Test for VLSI Circuits using LFSR. In *2023*

International Conference on Applied Intelligence and Sustainable Computing (ICAISC) (pp. 1-7). IEEE.

[2] Benabdenbi, M. (2023). *Contributions to the Test, Fault Tolerance and Approximate Computing of System on a Chip* (Doctoral dissertation, Université Grenoble Alpes).

[3] Bergamin, G. (2019). *Study of low power and radiation hard Design For Testability solutions for High Energy Physics applications* (Doctoral dissertation, Politecnico di Torino).

[4] Burger, E. S. (2020). Professional responsibility, legal malpractice, cybersecurity, and cyber-insurance in the COVID-19 era. . *Mary's J. on Legal Malpractice & Ethics*, *11*, 234.

[5] Chavan, A. (2022). Importance of identifying and establishing context boundaries while migrating from monolith to microservices. Journal of Engineering and Applied Sciences Technology, 4, E168. http://doi.org/10.47363/JEAST/2022(4)E168

[6] Chavan, A. (2022). Importance of identifying and establishing context boundaries while migrating from monolith to microservices. Helina. http://doi.org/10.47363/JEAST/2022(4)E168

[7] Grout, I., & A'ain, A. K. B. (2012). The Development of an Online Support Tool for the Teaching and Learning of the IEEE Standard 1500 for Embedded Core-based Integrated Circuits. *International Journal of Online Engineering*, *8*(4).

[8] Gupta, R., Gupta, P., & Singh, J. (2019). Security and Cryptography. In *Software Engineering for Embedded Systems* (pp. 501-547). Newnes.

[9] Hayat, K. (2022). Scalable 3D visualization via synchronous data hiding. *hal. archives-ouvertes. fr*.

[10] ISO, I. (2018). 26262 Road Vehicles-Function Safety-Part 5: Product development at the hardware level. *International Standardization Organization Std*.

[11] Jiang, W., Yu, H., Zhang, H., Shu, Y., Li, R., Chen, J., & Ha, Y. (2022). FODM: A framework for accurate online delay measurement supporting all timing paths in FPGA. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, *30*(4), 502-514.

[12] Kan, S., & Dworak, J. (2022). Towards Multipronged On-chip Memory and Data Protection From Verification to Design and Test.

[13] Karwa, K. (2023). AI-powered career coaching: Evaluating feedback tools for design students. Indian Journal of Economics & Business. https://www.ashwinanokha.com/ijeb-v22-4-2023.php

[14] Kellom, B., Monfalcone, M., & Pape, D. B. (2020). *Functional Safety Research Considerations for Heavy Vehicles* (No. DOT HS 812 922). United States. Department of Transportation. National Highway Traffic Safety Administration.

[15] Konneru, N. M. K. (2021). Integrating security into CI/CD pipelines: A DevSecOps approach with SAST, DAST, and SCA tools. *International Journal of Science and Research Archive*. Retrieved from https://ijsra.net/content/role-notification-scheduling-improving-patient

[16] Kumar, A. (2019). The convergence of predictive analytics in driving business intelligence and enhancing DevOps efficiency. International Journal of Computational Engineering and Management, 6(6), 118-142. Retrieved from https://ijcem.in/wp-content/uploads/THE-CONVERGENCE-OF-PREDICTIVE-ANALYTICS-IN-DRIVING-BUSINESS-INTELLIGENCE-AND-ENHANCING-DEVOPS-EFFICIENCY.pdf

[17] Kumar, B. (2020). *Effective Techniques for Post-silicon Validation and Debug* (Doctoral dissertation, Indian Institute of Technology Bombay Mumbai 400076, India).

[18] Madhura, R., & Shantiprasad, M. J. (2020). Implementation of Scan Logic and Pattern Generation for RTL Design. In *New Trends in Computational Vision and Bio-inspired Computing: Selected works presented at the ICCVBIC 2018, Coimbatore, India* (pp. 385-396). Cham: Springer International Publishing.

[19] Mohamed, S. (2023). *Protocol and Performance Optimization for Handling Heterogeneous Network Dynamics*. ProQuest Dissertations & Theses.

[20] Naveen Balaji, G., & Chenthur Pandian, S. (2019). Design of test pattern generator (TPG) by an optimized low power design for testability (DFT) for scan BIST circuits using transmission gates. *Cluster Computing*, *22*(Suppl 6), 15231-15244.

[21] Novák, O., Rozkovec, M., & Plíva, J. (2020). Decompressors using nonlinear codes. *Microprocessors and Microsystems*, *76*, 103076.

[22] Nyati, S. (2018). Revolutionizing LTL carrier operations: A comprehensive analysis of an algorithm-driven pickup and delivery dispatching solution. International Journal of Science and Research (IJSR), 7(2), 1659-1666. Retrieved from https://www.ijsr.net/getabstract.php?paperid=SR24203183637

[23] Nyati, S. (2018). Transforming telematics in fleet management: Innovations in asset tracking, efficiency, and communication. International Journal of Science and Research (IJSR), 7(10), 1804-1810. Retrieved from https://www.ijsr.net/getabstract.php?paperid=SR24203184230

[24] Posse, K. (2014). IEEE Std 1687-2014: IEEE Standard for Access and Control of Instrumentation Embedded within a Semiconductor Device. *IEEE Computer Society, Approval date*, *3*.

[25] Pradhan, M. (2020). *Studies on Diagnostic Coverage and X-Sensitivity in Logic Circuits* (Doctoral dissertation, Indian Statistical Institute-Kolkata).

[26] Puig-Rigall, J., Blanco-Prieto, M. J., Aydillo, C., Radulescu, A., Molero-Vilchez, D., Dreiss, C. A., & González-Gaitano, G. (2021). Poloxamine/D-α-Tocopheryl polyethylene glycol succinate (TPGS) mixed micelles and gels: Morphology, loading capacity and skin drug permeability. *Journal of molecular liquids*, *324*, 114930.

[27] Rajski, J., Tyszer, J., Kassab, M., & Mukherjee, N. (2004). Embedded deterministic test. *IEEE transactions on computer-aided design of integrated circuits and systems*, *23*(5), 776-792.

[28] Raju, R. K. (2017). Dynamic memory inference network for natural language inference. International Journal of Science and Research (IJSR), 6(2). https://www.ijsr.net/archive/v6i2/SR24926091431.pdf

[29] Reisinger, T. (2022). *Clock synchronization in a distributed hardware-in-the-loop testbed* (Doctoral dissertation, Technische Universität Wien).

[30] Repetti, T. J. (2023). *Designing, Implementing and Programming a Massively Multithreaded Spatial Accelerator Architecture*. Columbia University.

[31] Sardana, J. (2022). Scalable systems for healthcare communication: A design perspective. *International Journal of Science and Research Archive*. https://doi.org/10.30574/ijsra.2022.7.2.0253

[32] Sardana, J. (2022). The role of notification scheduling in improving patient outcomes. *International Journal of Science and Research Archive*. Retrieved from https://ijsra.net/content/role-notification-scheduling-improving-patient

[33] Singh, V. (2022). Integrating large language models with computer vision for enhanced image captioning: Combining LLMS with visual data to generate more accurate and context-rich image descriptions. Journal of Artificial Intelligence and Computer Vision, 1(E227). http://doi.org/10.47363/JAICC/2022(1)E227

[34] Sun, Y. (2021). *Novel test point insertion applications in Ibist* (Doctoral dissertation, Auburn University).

[35] Williams, T. W., & Daehn, W. (1989, January). Aliasing errors in multiple input signature analysis registers. In *Proceedings of the 1st European Test Conference* (pp. 338-339). IEEE Computer Society.

[36] Zhang, H., Jin, Y., Liu, Q., Zhao, Y., & Gao, Q. (2022). Intelligent monitoring method for tamping times during dynamic compaction construction using machine vision and pattern recognition. *Measurement*, *193*, 110835.