Boosting Verification Scalability via Structural Grouping and Semantic Partitioning of Properties

Rohit Dureja*, Jason Baumgartner†, Alexander Ivrii†, Robert Kanzelman† and Kristin Y. Rozier*
*Iowa State University †IBM Corporation

Abstract—From equivalence checking to functional verification to design-space exploration, industrial verification tasks entail checking a large number of properties on the same design. State-of-the-art tools typically solve all properties concurrently, or one-at-a-time. They do not optimally exploit subproblem sharing between properties, leaving an opportunity to save considerable verification resource via concurrent verification of properties with nearly identical cone of influence (COI). These high-affinity properties can be concurrently solved; the verification effort expended for one can be directly reused to accelerate the verification of the others, without hurting per-property verification resources through bloating COI size. We present a near-linear runtime algorithm for partitioning properties into provably high-affinity groups for concurrent solution. We also present an effective method to partition high-structural-affinity groups using semantic feedback, to yield an optimal multi-property localization abstraction solution. Experiments demonstrate substantial end-to-end verification speedups through these techniques, leveraging parallel solution of individual groups.

I. INTRODUCTION

The formal verification of a hardware and/or software design often mandates checking a large number of properties. For example, equivalence checking compares pairwise equality of each design output across two designs, and entails a distinct property per output. Functional verification checks designs against a large number of properties ranging from low-level assertions to high-level encompassing properties. Design-space exploration via model checking [16] verifies multiple properties against competing system designs differing in core capabilities or assumptions.

Each property has a distinct minimal cone of influence (COI), or fanin logic of the signals referenced in that property (Fig. 1a). Verification of a set of properties often entails exponential complexity with respect to the size of its collective COI. Concurrent verification of multiple properties may thus be significantly slower than solving these properties one-at-a-time, in that each property of the group may add unique fanin logic to the collective COI (Fig. 1b). Conversely, sometimes two or more properties share nearly-identical COIs (Fig. 1c). Concurrent verification of high-affinity properties may save considerable verification resource, as the effort expended for one can be directly reused for the others without significantly slowing the verification of any property within that group (e.g., reusing reachability clauses [5] [22], and abstractions in localization [1] across properties in a group).

Despite the prevalence of multi-property testbenches, little research has addressed the problem of optimal grouping or clustering of properties into high-affinity groups. Selective past work [7] [8] has experimentally demonstrated that ideal grouping may save substantial verification resource. However, no scalable online property grouping procedure has been provided; this potential was illustrated as a proof-of-concept using computationally-prohibitive offline grouping algorithms with undisclosed runtime. A significant need thus remains for an effective solution of determining which high-affinity properties should be concurrently solved. To ensure overall scalability, it is essential that such a property-partitioning solution be as close to linear runtime as possible with respect to the number of properties, otherwise the grouping effort itself may severely degrade overall verification resource comprising grouping plus subsequent verification of the identified groups.

Contributions: We present a near-linear runtime, fully-automated algorithm to partition properties into provably high-affinity groups based on structural COI similarity. COI support information is maintained as bitvectors [9], and grouping is performed in three configurable levels based on: identical COI, strongly-connected components (SCC) in the COI, and Hamming distance. The properties in each high-affinity group are verified concurrently; each group may be independently verified in parallel, using arbitrary solver algorithms. We also present an algorithm to semantically refine high-structural-affinity groups in a localization abstraction framework, offering the first optimized multi-property localization solution, to our knowledge. Our partitioning requires negligible resources even on the largest problems, while offering substantial verification speedups as demonstrated by extensive experiments.

A. Related Work

Much prior work has addressed methods to incrementally reuse information across multiple properties to accelerate specific algorithms. E.g., incremental SAT across proofs of different properties [20] [21], and reusing verification by-products like invariants [13] and interpolants [24], can accelerate the verification of high-affinity properties.

Methods to group properties based on high-level design descriptions extract similarity criteria from high-level information unavailable in low-level designs and benchmark formats
such as AIGs \[11\]. The framework of local and global proofs \[18\] has been used to derive a “debugging set” of properties to fix before verifying others, implying a property ordering but not a partitioning for minimal collective resource. LTL satisfiability checking has been used to establish logical dependencies between properties \[14\] to dynamically reduce verification resource; however, this work requires a quadratic number of resource-intensive comparisons.

The work most similar to ours is a property-clustering procedure based on COI similarity \[7, 8\]. While a similar goal, their solution requires a quadratic number of comparisons between properties, rendering it prohibitively expensive on large testbenches. Their experiments do not disclose grouping resource, only subsequent verification speedup. Moreover, this generic clustering approach requires the number of desired groups as an algorithmic parameter. This metric is impossible to predict in practice; it is far superior to allow affinity analysis to automatically determine the optimal number of groups.

II. PRELIMINARIES

A. Definitions

The logic design under verification is represented as a \textit{netlist}. A netlist contains a directed graph with \textit{gates} represented as vertices, and interconnections between vertices represented as edges. Every gate has an associated function: constants, primary inputs, combinational logic such as AND gates, and sequential logic such as \textit{registers}. Registers each have two associated gates that represent their next-state function, and their initial-value function. Semantically, the value of the register at time “0” equals the value of the initial-value function gate at time “0”, and the value of the register at time “i+1” equals that of the next-state function gate at time “i”.

Certain gates are labeled as \textit{properties}, formed through standard synthesis of the relevant property specification language. The \textit{fanin} of a property refers to the set of gates in the netlist which may be reached by traversing the netlist edges backward from the property gate. This fanin cone is called the \textit{cone-of-influence} (COI) of the property. The registers and inputs in the COI are called \textit{support variables}. The number of support variables in the COI is the COI size.

We say that a \textit{strongly connected component} (SCC) in the netlist is a set of interconnected gates such that there is a non-empty directed path between every pair of gates in the same SCC. In particular a primary input does not belong to any SCC, and in a well-formed SCC every directed cycle has at least one register because a netlist must be free of combinational cycles. The number of registers in a SCC is its \textit{weight}.

B. Cone-of-Influence Computation

Support variable information may be represented as an indexed array of Boolean values, or bitvector, per property. Fig. 2 gives a high-level procedure to compute a support bitvector for a property \(p\). Every support variable in the netlist \(N\) is indexed to an unique position in the bitvector, and \text{index}(v)\ returns that index for variable \(v\). The function \text{fanin}(p)\ recursively computes the fanin structure, or COI, of the gate corresponding to property \(p\). If a support variable \(v\) is in the fanin of property \(p\), then the \text{index}(v)\th bit is set to “1” in the bitvector; otherwise, is set to “0”. The \text{length} of such a bitvector is equal to the total number of support variables in the netlist, and all bitvectors have the same length. The COI size of the property is the number of bits set to “1”, and can be computed using fast population counting algorithms \[31\].

The bitvectors can be \textit{packed} by representing every SCC as a single weighted support variable; SCC bits have weight equal to the SCC weight, while others have unit-weight. The COI size of the property equals the weighted sum of the bits set to “1”. Note that the choice of whether or not to represent SCCs as a single bit does not affect the resulting support size. Unless stated otherwise, a “support bitvector” is assumed packed.

Practically, it is far too computationally expensive to walk the fanin cone of every property independently. Instead, the netlist may be traversed once in a topological manner, computing intermediate support bitvectors for internal gates \[9\]. E.g., for an AND gate \(a_1\) with incoming edges \(i_1\) and \(i_2\), the intermediate bitvector for \(a_1\) is simply the disjunction over the bitvectors for \(i_1\) and \(i_2\). For more details on support bitvector computation and optimizations, we refer the reader to \[9\].

C. Property Affinity

The \textit{unpacked} bitvectors for every property can be analyzed to determine \textit{affinity} among the properties. We use “Hamming distance” as an affinity measure; high-affinity properties have nearly-identical bitvectors. The affinity between properties \(p_1\) and \(p_2\) with unpacked bitvectors \(bv_1\) and \(bv_2\) is:

\[
0 \leq \text{affinity}(p_1, p_2) = 1 - \frac{\text{hamming}(bv_1, bv_2)}{\text{length}(bv_1)} \leq 1.0
\]

where \text{hamming}(bv_1, bv_2)\ is the Hamming distance between the unpacked bitvectors, and \text{length}(bv_1)\ is the number of support variables in the netlist (identical for every bitvector). Let \(V_1\) and \(V_2\) be the set of support variables in the COI of \(p_1\) and \(p_2\), respectively. Note that \text{hamming}(bv_1, bv_2)\ equals (\(|V_1 \cup V_2| - |V_1 \cap V_2|\)) and \text{length}(bv_1)\ equals \(|V_1 \cup V_2|\).

D. Group Center and Grouping Quality

A property \(p\) is selected in a group \(g\) that represents the group’s \textit{center}, or \textit{representative} property, and is denoted as \(g^*\). The \textit{quality} of a group \(g\), denoted \(Q(g)\), is the minimum affinity between any pair of properties in \(g\), i.e.,

\[
Q(g) = \min \{ \text{affinity}(p_i, p_j) \mid \forall p_i, p_j \in g \}
\]

A quality of \(t\) implies that \textit{unpacked} bitvectors, of length \(l\), for properties in a group have a maximum Hamming distance of
(1 − t) * l. Our grouping algorithms guarantee that the quality of every group will be greater than a specified threshold.

E. Localization Abstraction

The proof or counterexample for a property often only depends on a small subset of its COI logic. Localization abstraction \([1, 10, 25, 26]\) is a powerful method aimed at reducing netlist size by removing irrelevant logic, transforming irrelevant gates to unconstrained primary input variables via cutpoint insertion. Since cutpoints can simulate the behavior of the original gates and more, the localized netlist over-approximates the behavior of the original netlist. Abstraction refinement is used to eliminate cutpoints which are deemed responsible for any spurious counterexamples, effectively re-introducing previously-eliminated logic. Ultimately, the abstract netlist is passed to a proof engine. It is desirable that the abstract netlist be as small as possible to enable more-efficient proofs, while being immune to spurious counterexamples.

III. Structural Grouping of Properties

Practical industrial verification tasks often entail hundreds of thousands of support variables, and tens of thousands of properties. The need for scalability obviates straightforward approaches, such as pairwise-comparing each property to check for affinity. We use support bitvectors for a set of properties, and partition them into high-affinity property groups. Our affinity-based algorithm performs grouping in three configurable levels based on: identical bitvectors (level-1), weights of large SCCs in support (level-2), and Hamming distance between bitvectors (level-3). The underlying intuition is that properties with similar bitvectors, measured in terms of a distance metric like Hamming distance, have high structural affinity and can be most efficiently verified as one concurrent multi-property verification task. To ensure overall scalability, each level runs in as close to linear runtime as possible with respect to the number of properties, otherwise the grouping effort itself may severely degrade overall verification resource comprising grouping plus verification of the identified groups.

Fig. 3 shows our leveled structural grouping algorithm. It takes as input the properties P, netlist N, and desired grouping level l. Additionally, an affinity threshold t controls the quality of groups formed. Each property is initially assigned its own distinct group, i.e., each group contains only one property. Upon termination, properties in a group are checked concurrently using a verification algorithm portfolio, and different groups are verified independently.

A. Level-1 Grouping – Identical COI

The procedure to perform property grouping based on identical support bitvectors is demonstrated in Fig. 4. The procedure takes an initial property grouping as input, and then merges groups that have identical support bitvectors. g∗ denotes the representative property in a group, i.e., g* is the center. The choice of g* is trivial at level-1 because every group contains only one property. Next, the support bitvector for the center property in the group is hashed to an integer value. The choice of the hash function is implementation-dependent. We use Murmur3 \([2]\) to hash bitvectors as being very fast and accurate with minimal collisions, however, other functions can also be used. Groups for which the bitvector hashes to the same integer value, and further which have identical bitvectors, are then merged. Any property in the merged group can be chosen as the new center property without affecting subsequent results.

**Theorem 1.** Level-1 grouping generates high-affinity property groups G such that ∀g ∈ G : Q(g) = 1.0.

**Proof.** Initially, every group contains one property. Properties with identical bitvectors, i.e., affinity = 1.0, are grouped.

While scalable (near-linear runtime) and able to group properties with 100% affinity, in practice it is desirable to perform additional grouping of properties which have a small tolerable Hamming distance yet are still high-affinity. Again, we stress that a simple procedure of pairwise comparison to check whether properties are within a small tolerance is prohibitively slow in practice, rendering prior techniques as \([7, 8]\) unusable in practice. The following algorithms solve this goal of high-affinity group merging, with high scalability and guaranteed grouping quality.

B. Level-2 Grouping – Heavy-weight SCCs in COI

Many practical netlists contain at least one very large SCC, comprising the majority of its registers. For such netlists,
Given affinity $t$, level-2 grouping generates property groups $G$ such that $\forall g \in G : Q(g) \geq t$.

Proof. (Sketch) Let $n$ be the number of support variables. Properties with identical heavy-weight SCCs in their COI that contain $t\%$ of the variables have the same $t \times n$ bits set to “1” in their unpacked bitvectors, implying a maximum Hamming distance of $(1 - t) \times n$ or minimum affinity of $t$. □

Properties sharing a small number of common large SCCs may thus be adequately high-affinity to group based solely upon analysis of these SCCs, without needing to consider a potentially very large number of non-SCC support variables or smaller SCCs. In contrast, storing every full bitvector in a trie may become computationally expensive and serve little benefit. Since the subsequent level-3 grouping does not contain the same heavy-weight SCCs in their COI can often be grouped together as having high affinity. Fig. 5 demonstrates the procedure to perform property grouping based on heavy-weight SCCs. The procedure takes as input an affinity threshold $t$. For every group $g$, we find all SCCs in the COI of the center property $p = g^*$, with weight at least $w$. We use Tarjan’s algorithm to find SCCs in the COI of property $p$ in linear runtime. Practically, it is very expensive to find SCCs in the COI of every property independently. Instead, all SCCs are computed once for netlist $N$ along with the linear traversal to compute support bitvectors for properties $[23]$. If the cumulative SCC weight is at least $t$ times the number of support variables in netlist $N$, this set of SCCs is inserted into a prefix tree or trie (for fast ~linear time lookup and prefix matching). A hash table may be used, at the expense of possibly-increased memory footprint. If the trie already contains this set of SCCs, albeit for another group $h$, the two groups are merged. Any property in the merged group can be chosen as the new center property without affecting subsequent results.

### Theorem 2

Given affinity $t$, level-2 grouping generates property groups $G$ such that $\forall g \in G : Q(g) \geq t$.

Proof. (Sketch) Let $n$ be the number of support variables. Properties with identical heavy-weight SCCs in their COI that contain $t\%$ of the variables have the same $t \times n$ bits set to “1” in their unpacked bitvectors, implying a maximum Hamming distance of $(1 - t) \times n$ or minimum affinity of $t$. □

Properties sharing a small number of common large SCCs may thus be adequately high-affinity to group based solely upon analysis of these SCCs, without needing to consider a potentially very large number of non-SCC support variables or smaller SCCs. In contrast, storing every full bitvector in a trie may become computationally expensive and serve little benefit. Since the subsequent level-3 grouping does not contain the same heavy-weight SCCs in their COI can often be grouped together as having high affinity.
is at least $\hat{t} = 1 - \left\lfloor (1 - t) \cdot n \right\rfloor \div n$ ($\left\lfloor x \right\rfloor$ is the nearest integer function); a map $m$ is maintained that maps every $n$-bit number $(0, 1, \ldots, 2^n - 1)$ to the allotted cluster center $(1, \ldots, k)$. For a fixed value of $n$, the number of clusters $k$ can be increased one-by-one until quality of each is at least $t$, i.e. the maximum Hamming distance allowed per $n$-bit segment in a cluster is $(1 - \hat{t}) \cdot n$. E.g. for $n = 32$ and $t = 0.95$, the maximum hamming distance is $(1 - 0.95) \cdot 32 = 1.6 \approx 2$ for which $\hat{t} = 0.9375$. It is important to note that the initialization step involving clustering does not hinder scalability because:

1. The value of $n$ is typically less than the maximum CPU word size that allows fast single-cycle Hamming distance computation between two numbers (xor); clustering with $t = 0.9$ for $n = 16$ and $t = 0.9$ takes <1s and <1min, resp.
2. The map can be computed once offline, and reused in all future runs of the algorithm (e.g. embedded into a verification tool) for various ranges of threshold $t$.
3. For online computation, an approximate linear-time algorithm, like Gonzalez [19], can be used on $S$ that may only contain $n$-bit numbers appearing in bitvectors $BV$.

In the clustering step, every unpacked bitvector $bv$ is read in $n$-bit segments to generate a piecewise-mapped bitvector $mbv$ using map $m$. Bitvectors for which the corresponding mapped bitvectors hash to the same value are put in the same cluster.

**Theorem 3.** Given unpacked bitvectors $BV$, affinity $t$, and word size $n$, bitvector_cluster returns clusters $C$ such that $\forall c \in C : Q(c) \geq \hat{t}$, where $\hat{t} = 1 - \left\lfloor (1 - t) \cdot n \right\rfloor \div n$.

**Proof.** (Sketch) Assume generate_map() creates clusters from $n$-bit numbers in $S$ such that minimum affinity between numbers in each cluster is $t$, implying a Hamming distance of $(1 - t) \cdot n$. Let each unpacked bitvector $bv \in BV$ contain $num$ $n$-bit segments, i.e., $length(bv) = n \cdot num$. If two mapped bitvectors hash to the same value, then every $i^{th}$ $n$-bit segment in the two original unpacked bitvectors is at a maximum distance of $(1 - \hat{t}) \cdot n$. Therefore, the maximum distance between the two unpacked bitvectors is $(1 - \hat{t}) \cdot n \cdot num$ or $(1 - \hat{t}) \cdot length(bv)$, implying a minimum affinity of $\hat{t}$.

Fig. [7] demonstrates the procedure to perform property grouping based on Hamming distance using the bitvector clustering algorithm of Fig. [6]. The algorithm generates a map $m$ of $n$-bit numbers to cluster centers $1, \ldots, k$ as an initialization step. The center property unpacked bitvector for every group is read per $n$-bit segment, to generate a mapped bitvector using map $m$. The mapped bitvector is hashed to an integer value. The groups for which the center property mapped bitvectors hash to the same value, and further which have identical mapped bitvectors, are immediately grouped.

**Theorem 4.** Given affinity $t$ and word size $n$, level-3 grouping generates property groups $G$ such that $\forall g \in G : Q(g) \geq 2 \cdot t + \hat{t} - 2$, where $\hat{t} = 1 - \left\lfloor (1 - t) \cdot n \right\rfloor \div n$.

**Proof.** (Sketch) The proof follows from triangle inequality of Hamming distance. Let $m$ be the length of unpacked bitvectors. For groups $g_1$ and $g_2$, if center property bitvectors hash to the same value then the bitvectors are at a distance of at most $(1 - \hat{t}) \cdot m$ (Theorem 5). Properties within groups $g_1$ and $g_2$ are at a maximum distance of $(1 - t) \cdot m$ (Theorems 1 & 2) from their respective center properties. Therefore, maximum distance between a property in $g_1$ and another property in $g_2$ is $2 \cdot (1 - t) \cdot m + (1 - \hat{t}) \cdot m, or (1 - (2 \cdot t + \hat{t} - 2)) \cdot m$, implying a minimum affinity of $2 \cdot t + \hat{t} - 2$.

When $\hat{t} = t$, level-3 returns groups with $Q(g) \geq 3 \cdot t - 2$. Despite its provable threshold, there is some asymmetry in this approach, in that two fairly-high-affinity bitvectors which differ too much in a single segment will not be merged, whereas if the difference was small per-segment with multiple segments differentiated, they may be merged, albeit respecting the quality bound. The highly scalable analysis can be repeated if higher precision and symmetry is desired. This can be done either as-is on the entire netlist under different permutations or segment-partitioning of bitvector indices (i.e., by varying the starting index of the first $n$-bit segment in the bitvector), or on individual (sets of) groups obtained from the prior run. Since re-running on a subset of properties implies a smaller cone-of-influence, bitvectors can be compacted for faster runtime to only include support variables in the COI of any considered property, and this indexing will differ from the prior run over a larger set of properties. Moreover, support variables present in the COI of every property can be completely projected out of the bitvectors to offer further compaction and speedup.

**IV. SEMANTIC REFINEMENT OF PROPERTY GROUPS**

It is desirable that the netlist generated by localization abstraction be as small as possible to enable efficient proofs. Localization cutpoints are property-specific, hence concurrent localization of properties with disjoint COIs - or even similar COIs - may yield significantly larger netlists which are less
localization (Group \( g \), Netlist \( N \), Limit \( n \), Threshold \( t \))

1. Netlist \( L \) # localized netlist
2. \( L = \text{initial_abstraction}(g) \) # add gates for every property
3. unsigned \( k = 0 \) # bmc depth
4. bool stop = 0 # some properties fail at depth \( k \)
5. while not stop : # loop until all properties pass at depth \( k \)
6. stop = 1
7. Gates \( c = \{ \} \) # cutpoints to refine in \( L \), initially empty
8. for each Property \( p \in g : \)
9. \( \text{Result} = \text{run_bmc}(L, p, k) \) # run bmc with depth \( k \)
10. if \( \text{Result == unsat} : \) continue # property passes
11. # check counterexample returned by bmc
12. if cex not spurious : report_solved(p, cex), continue
13. stop = 0 # property fails
14. Gates \( d = \text{cutpoints_to_refine}(c, c \cup d) \) # add to support bitvector
15. # at least one property passes at depth \( k \)
16. if not stop : \text{refine_abstraction}(L, c), unchanged = 0
17. else unchanged + = 1 # no change in abstraction
18. # check if netlist unchanged for last \( n \) bmc steps
19. if unchanged < \( n \) : \( k = k + 1 \) goto line 4 # increment depth
20. else Groups \( \hat{G} = \text{structural_grouping}(g, L, 3, t) \)
21. # run proof engine for each group in \( \hat{G} \) with netlist \( L \)

Fig. 8. Localization to partition a group \( g \) of high-affinity properties. BMC is run for increasing depth until there is no change in the localized netlist, after which partitioning is attempted to split \( g \) into subgroups \( \hat{G} \).

scalable to verify. Our structural property grouping procedure ensures that only high-affinity properties in a group will be localized concurrently, which helps ensure smaller multiproperty abstractions. However, it might be the case that a cutpoint is refined for one property in a high-affinity group, whereas that refinement may be unnecessary for another property in the group. As a result, properties in a high-affinity group without localization cutpoints may have vastly different COI in the localized netlist. Therefore, partitioning the group obtained from Fig. 3 into high-affinity localized subgroups based upon localization decisions can improve overall verification scalability.

A. Generating Support Bitvectors

Various techniques have been proposed [1, 10, 25] to guide the abstraction-refinement process of localization. Most-state-of-the-art localization implementations use SAT-based bounded model checking (BMC) [4] to select the localized netlist upon which an unbounded proof is attempted. In our implementation we run BMC iteratively until there is no change in the localized netlist. Fig. 8 shows our localization abstraction framework that supports high-affinity group partitioning. We start with a localized netlist only containing property gates. For a given BMC depth \( k \), we iterate over properties in group \( g \) to eliminate all spurious counterexamples of length \( k \). Cutpoints deemed necessary to refine for a property \( p \) are collected (line 13). If a cutpoint is also a support variable, it is then added to the support bitvector maintained for property \( p \) (line 14). The abstraction is then refined using the collected cutpoints, and BMC is run again at depth \( k \).

When all properties hold for the abstract model at depth \( k \), BMC is run again with depth \( k + 1 \). The repeated BMC runs add new cutpoints to the support bitvector for every property, which in turn can be used to partition group \( g \) into high-affinity subgroups with respect to the localized netlist. Various strategies may be used to decide when to terminate BMC: an upper-bound on BMC depth or runtime can be used. In our framework, we prefer increasing BMC depth until there is no change in the localized netlist for \( n \) consecutive steps (lines 17–18). The value of \( n \) can be varied to increase confidence in the abstracted model such that it is immune to spurious counterexamples.

B. Group Partitioning

Once BMC converges, group \( g \) is then partitioned into subgroups \( \hat{G} \) based on support bitvector information. Note that the problem is analogous to grouping of properties in \( g \) with respect to the localized netlist. Therefore, we use the property grouping procedure of Fig. 3 to generate high-affinity property groups for overall scalability. The properties in each subgroup are then passed to a proof engine for verification with respect to each COI-reduced localized subgroup’s netlist.

V. EXPERIMENTAL RESULTS

We experimentally analyze our grouping procedure on end-to-end verification scalability. Our grouping procedure is implemented within Rulebase: Sixthsense Edition [27]. All experiments were run on Linux machines, with 32GB memory. Time reported is `cpu` time. We refer to different grouping levels as \( L_1 \), \( L_2 \), and \( L_3 \).

A. Benchmarks from HWMCC

We evaluate 48 benchmarks from HWMCC that contain more than 100 safety properties (Fig. 9a). These are obtained by simplifying all the benchmarks by standard logic synthesis (similar to \&dc2 in ABC [6]) to solve easy properties, and disjunctive decomposition to fragment each OR-gate property into a sub-property of its literals. Each property, or property group, is solved using a portfolio comprising BMC [4], IC3 [5], and localization (LOC) without semantic partitioning. Each can process multiple properties: IC3 and BMC in a time-sharing manner, and LOC concurrently abstracting a set of properties which are solved using IC3.

a) Property Grouping: Support bitvector computation is fast, and takes less than five seconds on the largest benchmark. The ideal threshold is benchmark- and solver-specific. Given the exponential penalty of grouping lower-affinity properties vs. linear penalty of splitting higher-affinity properties (offset by parallel solving), we find it best to err to the latter using a higher affinity \( t = 0.9 \). \( L_3 \) is done using 16-bit words and

1. Off-the-shelf clustering is more applicable here than on the original netlist if desired, because: (1) the localized netlist and support bitvectors are often immensely smaller than the original netlist; (2) the number of properties per structural group being localized is often smaller than the number of overall netlist properties. However, there is no guarantee of either of these points.

2. Detailed results available at http://temporallogic.org/research/FMCAD19


(b) Grouping Time (ms)

![Graph](image)

Fig. 9. Grouping on 48 HWMCC benchmarks with more than 100 properties, and maximum 50 properties/group. Level-1 grouping quality is 1.0.

(c) Reduction Ratio

(d) Grouping Quality

![Graph](image)

Fig. 10. End-to-end verification with grouping vs. portfolio which (a) checks properties one-at-a-time, and (b) check all properties together. Points below diagonal are in favor of verification with grouped properties.

\[ i = 0.875, \text{i.e., maximum distance of 2 between words} \]

(Fig. 9). Initially each property is assigned to its own distinct group. The grouping takes less than 10ms for all benchmarks (Fig. 9b). The group count reduction ratio for each level with respect to the preceding level, and overall reduction ratio, i.e., number of groups relative to preceding level, is shown in Fig. 9c. \( L_2 \) merges properties for 13 benchmarks: <0.5 ratio for 8 benchmarks, and is critical to the performance of \( L_3 \). Without \( L_1 \) and \( L_2 \), not all properties merged by \( L_2 \) are merged by \( L_3 \) due to inherent asymmetry, and \( L_3 \) can merge the same properties as \( L_1 \), albeit, with small runtime penalty. Therefore, the leveling order is crucial and gives tighter control on group affinity. Fig. 9d shows the minimum quality of all non-singleton groups in a benchmark.

(b) End-to-end Verification: We compare the runtime of checking each property one-by-one vs. checking property groups in Fig. 10a. Verification with structural grouping is up to 400× (median 4.3×) faster. A fairer comparison of the runtime of checking all properties together vs. checking property groups is shown in Fig. 10b. Grouped verification is up to 72× (median 3.5×) faster. Table I shows benchmarks for which checking all properties together is faster. LOC solves very few properties for these benchmarks, whereas, BMC/IC3 quickly verify all properties together: 145 properties in 6s380 are falsified by BMC in a few unrollings and remaining proved by LOC, while all properties are proved by LOC or IC3 for other benchmarks. The benchmarks in Table I have properties where a large majority are either all falsified, or proved. The advantage of checking high-affinity groups is outweighed by the added cost of repeating BMC/IC3 across groups for these benchmarks, which could be adjusted for using a lower affinity threshold. However, grouping advantage is apparent for benchmarks in which no single algorithm solves all properties, and properties have different verification outcomes.

c) Localization Abstraction: We select 24 benchmarks having at least 50 properties solved by LOC. Properties not solved by LOC are not considered. Fig. 11a shows the impact of high and low affinity grouping on the performance of LOC. If high-affinity structural grouping returns \( N \) groups, low-affinity grouping is done by sorting properties by COI size, and partitioning into equally-sized \( N \) groups. Fig. 11a compares verification of high-affinity grouped properties and one-by-one checking of each property with LOC; verification is up to 30× (median 2.9×) faster. On the other hand, low-affinity groups often degrade LOC performance compared to one-by-one checking. Fig. 11b compares high and low affinity group verification with LOC. Five benchmarks have comparable performance due to grouping of large number of properties into very few groups. Nevertheless, high-affinity verification is always faster: up to 3.7× (median 2.5×).

d) Semantic Partitioning: LOC generates a localized netlist using BMC for every property group which is then checked by a proof engine. If the localization is sufficient,
the proof engine may prove all properties in a single run. Otherwise, it generates a possibly-spurious counterexample. Table II shows benchmarks in which some non-singleton groups are proved by LOC in a single proof-engine run. We perform semantic partitioning on these groups. ‘Total’ shows the #Groups generated by structural grouping for #Props, whereas, ‘Single Run’ shows the #Groups and #Props solved by one proof engine run after generating a sufficient localized netlist. All groups are solved by LOC one-by-one. As is evident, semantic partitioning boosts the performance of LOC for hard problems (in bold). However, there is a marginal slowdown for easy problems due to the overhead of restarting the proof engine on semantically partitioned subgroups.

e) Lossy Grouping: Lastly, we compare the grouping loss using our procedure with hierarchical clustering (HC). We measure loss as #properties assigned a group by HC but not our procedure (maximum 50 properties/group). Table III summarizes results for five representative benchmarks. HC always takes more grouping time. There is no loss in benchmarks for which both methods return very few groups (e.g., 6s405). Verification with fewer groups from HC is faster (e.g., 6s381) when our procedure has higher loss. This loss may be due to prefix mismatch, and asymmetry in L3, which can be mitigated by using techniques in Sec. III-C. In most cases, HC gives fewer groups which may result in less verification time, but HC grouping resource results in an end-to-end runtime degradation vs. our approach. It is clear that HC gives tighter groups but overall verification resource is dominated by the time it takes to perform grouping.

### B. Proprietary Designs

Post-silicon observability solutions often leverage monitoring logic instrumented throughout a hardware design. This debug bus logic monitors a configurable set of internal signals in real-time, non-intrusively while the chip is functionally running. Debug bus verification entails a large number of properties (often one per monitor point), within very large design components - sometimes entire chips [17]. Localization is the dominant method to verify debug bus designs as they often contain >10M gates [17]. Note that concurrent verification of all properties is completely intractable. Table IV summarizes our results. ‘One-by-one’ shows verification time by localizing one property at a time, whereas, ‘Grouped’ represents concurrent localization of properties in a high-affinity group. All designs benefit from high-affinity group verification, and the speedup is clearly evident for large designs (in bold).

### VI. CONCLUSIONS AND FUTURE WORK

Scalable property grouping is a hard problem. Existing approaches are either syntax-based [11], or resource intensive [7]. The need for scalability cannot be over-stated; traditional grouping algorithms require at least quadratic runtime vs. number of properties, and are prohibitively slow–adding to and easily outweighing the benefit they bring to the verification process. We present a 2-step grouping strategy: structural grouping followed by semantic partitioning, that offers massive end-to-end verification speedup. Experiments demonstrate the usefulness of our method on several verification tasks: structural grouping is trivially fast regardless of subsequent verification engines, and semantic partitioning accelerates difficult localization problems. We advance state-of-the-art in localization by providing an optimal multi-property solution.

Future work includes improved ordering and compaction of support bitvector bits to improve performance, e.g., support variables present in every property can be projected out of the bitvectors. Dynamic trie matching that discounts differences in very small SCCs in COI for properties, may improve level-2 grouping. Extending level-3 grouping to work with packed bitvectors may speed up grouping: large SCCs for which any distinction exceeds threshold require identical valuations in grouping, and smaller SCCs are either unpacked to multiple bits or treated with finer-grained map. Clever data structures, such as MA-FSA [12], and branch-and-bound traversal [30] can search for fairly-high-affinity bitvectors that differ in only a few n-bit segments, thereby reducing level-3 asymmetry. Extending semantic partitioning to cases where refinement occurs during a proof engine run is a promising research direction. We plan to investigate how semantic information from BMC and IC3 can be used to perform property grouping.

---

**TABLE II**

<table>
<thead>
<tr>
<th>Name</th>
<th>Total</th>
<th>Single Run</th>
<th>Verification Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#G</td>
<td>#P</td>
<td>Dis.</td>
</tr>
<tr>
<td>6s384</td>
<td>2</td>
<td>51</td>
<td>1</td>
</tr>
<tr>
<td>6s344</td>
<td>12</td>
<td>247</td>
<td>3</td>
</tr>
<tr>
<td>6s405</td>
<td>13</td>
<td>593</td>
<td>3</td>
</tr>
<tr>
<td>6s410</td>
<td>15</td>
<td>735</td>
<td>4</td>
</tr>
<tr>
<td>6s310</td>
<td>15</td>
<td>186</td>
<td>5</td>
</tr>
<tr>
<td>6s391</td>
<td>30</td>
<td>144</td>
<td>9</td>
</tr>
<tr>
<td>6s332</td>
<td>77</td>
<td>163</td>
<td>16</td>
</tr>
</tbody>
</table>

**TABLE III**

<table>
<thead>
<tr>
<th>Name</th>
<th>#P</th>
<th>Our Procedure</th>
<th>Hierarchical</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#G</td>
<td>G.Time</td>
<td>#G</td>
</tr>
<tr>
<td>6s405</td>
<td>593</td>
<td>2.16ms</td>
<td>13</td>
</tr>
<tr>
<td>6s381</td>
<td>1506</td>
<td>5.93ms</td>
<td>76</td>
</tr>
<tr>
<td>6s361</td>
<td>2653</td>
<td>12.71ms</td>
<td>62</td>
</tr>
<tr>
<td>6s117</td>
<td>8063</td>
<td>25.53ms</td>
<td>165</td>
</tr>
<tr>
<td>6s114</td>
<td>30628</td>
<td>2.65h</td>
<td>873</td>
</tr>
</tbody>
</table>

*Not simplified by logic synthesis*

---

**TABLE IV**

<table>
<thead>
<tr>
<th>ID</th>
<th>#P</th>
<th>#G</th>
<th>G.Time (ms)</th>
<th>Verification Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>36</td>
<td>9</td>
<td>0.48</td>
<td>32.69s 19.27s 1.70x</td>
</tr>
<tr>
<td>2</td>
<td>45</td>
<td>3</td>
<td>0.49</td>
<td>26.24s 12.63s 2.08x</td>
</tr>
<tr>
<td>3</td>
<td>56</td>
<td>5</td>
<td>0.94</td>
<td>11.9s   6.34s 1.88x</td>
</tr>
<tr>
<td>4</td>
<td>78</td>
<td>6</td>
<td>3.87</td>
<td>0.21s   0.14s 1.40x</td>
</tr>
<tr>
<td>5</td>
<td>148</td>
<td>4</td>
<td>6.85</td>
<td>95.88s  22.68s 4.23x</td>
</tr>
<tr>
<td>6</td>
<td>224</td>
<td>6</td>
<td>0.74</td>
<td>65.52s  19.65s 3.34x</td>
</tr>
<tr>
<td>7</td>
<td>1506</td>
<td>53</td>
<td>9.16</td>
<td>0.93h   0.21h 4.32x</td>
</tr>
<tr>
<td>8</td>
<td>9371</td>
<td>1027</td>
<td>137.72</td>
<td>52.65s  11.89s 4.43x</td>
</tr>
<tr>
<td>9</td>
<td>11035</td>
<td>1238</td>
<td>146.32</td>
<td>7.94h   2.81h 2.82x</td>
</tr>
</tbody>
</table>

**TABLE V**

<table>
<thead>
<tr>
<th>Name</th>
<th>Total</th>
<th>Single Run</th>
<th>Verification Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rows</td>
<td>G.Time</td>
<td>Rows</td>
</tr>
<tr>
<td>6s384</td>
<td>2</td>
<td>22.65s</td>
<td>2</td>
</tr>
<tr>
<td>6s344</td>
<td>12</td>
<td>2.04s</td>
<td>12</td>
</tr>
<tr>
<td>6s405</td>
<td>13</td>
<td>0.28s</td>
<td>13</td>
</tr>
<tr>
<td>6s410</td>
<td>15</td>
<td>0.18s</td>
<td>15</td>
</tr>
<tr>
<td>6s310</td>
<td>15</td>
<td>82.13s</td>
<td>15</td>
</tr>
<tr>
<td>6s391</td>
<td>30</td>
<td>25.61s</td>
<td>30</td>
</tr>
<tr>
<td>6s332</td>
<td>77</td>
<td>1.21h</td>
<td>77</td>
</tr>
</tbody>
</table>

*Not simplified by logic synthesis*
ACKNOWLEDGMENTS

We thank the anonymous reviewers for their valuable comments. We thank Gianpiero Cabodi for an insightful discussion on multi-property verification, and Brad Bingham for feedback on early drafts of the paper. This work is partially supported by NSF CAREER Award CNS-1664356.

REFERENCES


