

Checking and Enforcing Robustness against TSO

Ahmed Bouajjani¹, Egor Derevenets^{2,3}, and Roland Meyer³

¹LIAFA, University Paris 7 ²Fraunhofer ITWM ³University of Kaiserslautern

Abstract. We present algorithms for checking and enforcing robustness of concurrent programs against the Total Store Ordering (TSO) memory model. A program is robust if all its TSO computations correspond to computations under the Sequential Consistency (SC) semantics.

We provide a complete characterization of non-robustness in terms of so-called attacks: a restricted form of (harmful) out-of-program-order executions. Then, we show that detecting attacks can be parallelized, and can be solved using state reachability queries under SC semantics in a suitably instrumented program obtained by a linear size source-to-source translation. Importantly, the construction is valid for an arbitrary number of addresses and an arbitrary number of parallel threads, and it is independent from the data domain and from the size of store buffers in the TSO semantics. In particular, when the data domain is finite and the number of addresses is fixed, we obtain decidability and complexity results for robustness, even for an arbitrary number of threads.

As a second contribution, we provide an algorithm for computing an optimal set of fences that enforce robustness. We consider two criteria of optimality: minimization of program size and maximization of its performance. The algorithms we define are implemented, and we successfully applied them to analyzing and correcting several concurrent algorithms.

1 Introduction

Sequential consistency (SC) [21] is a natural shared-memory model where the actions of different threads are interleaved while the program order between actions of each thread is preserved. However, for performance reasons, modern multiprocessors implement weaker memory models relaxing program order. For instance, the common store-to-load relaxation, which allows loads to overtake stores, reflects the use of *store buffers*. It is actually the main feature of the TSO (Total Store Ordering) model adopted, e.g., in x86 machines [27].

Nonetheless, most programmers usually assume that memory accesses are performed according to SC where all shared-memory accesses are instantaneous and atomic. This assumption is actually safe for *data-race-free* programs [3], but in many situations data-race-freedom does not apply. This is, for instance, the case of programs implementing synchronization operations, concurrency libraries, and other performance-critical system services employing lock-free synchronization. In most cases, programmers design programs that are *robust* against relaxations, i.e., for which relaxations do not introduce behaviors that are not allowed under SC. Memory fences must be included appropriately

in programs in order to prevent non-SC behaviors. Getting such programs right is a notoriously difficult and error-prone task. Therefore, important issues in this context are (1) checking robustness of a program against relaxations of a memory model, and (2) identifying a set of program locations where it is *necessary* to insert fences to ensure robustness.

In this paper we address these two issues in the case of TSO. We consider a general setting without fixed bounds on the shared memory size, nor on the size of the store buffers in the TSO semantics, nor on the number of threads. This allows us to reason about robustness of general algorithms without assuming any fixed values for these parameters that depend on the actual machine's implementation. Moreover, we tackle these issues for general programs, independently from the domain of data they manipulate.

Robustness against memory models has been addressed first by Burckhardt and Musuvathi in [9] (actually, for TSO only), and subsequently by Burnim et al. in [10]. Alglave and Maranget developed a general framework for reasoning about robustness against memory models in [4,5] (where the term *stability* is used instead of robustness). Roughly, these works are based on characterizing robustness in terms of acyclicity of a suitable happens-before relation. In that, they follow the spirit of Shasha and Snir [28] who introduced a notion of *trace* that captures the control and data dependencies between events of an SC computation, and established that computations that are not SC have a happens-before relation that contains a cycle. We adopt here the same notion of robustness, i.e., a program is (trace-)robust if each of its TSO computations has the same trace as some of its SC computations.

From an algorithmic point of view, the existing works mentioned above *do not provide decision procedures* for robustness. [9,10] provide testing procedures based on enumerating TSO runs and checking that they do not produce happens-before cycles. Clearly, while these procedures can establish non-robustness, they can never prove a program robust. On the other hand, [5] provides a sound over-approximate static analysis that allows to prove robustness, but may also inaccurately conclude to non-robustness and insert fences unnecessarily. We are interested here in developing an approach that allows for precise checking of trace-robustness, and for optimal fence insertion (in a sense defined later).

In our previous work [8], trace-robustness against TSO has been proved to be decidable and PSPACE-complete, even for unbounded store buffers, in the case of a fixed number of threads and assuming a fixed number of shared variables, ranging over a finite data domain. The method that shows this decidability and complexity result does not provide a practical algorithm: it is based on a non-deterministic, bounded enumeration of computations. Moreover, it does not carry over to the general setting we consider here. Therefore, in this paper we propose a novel approach to checking robustness that is fundamentally different from [8]. We provide a general, source-to-source reduction of the trace-robustness problem against TSO to the *state reachability problem under SC semantics*. In other words, we show that trace-robustness is not more expensive than SC state reachability, which is the unavoidable problem to be solved by any precise

decision algorithm for concurrent programs. This is the *key contribution of the paper* from which we derive other results, such as decidability results in particular cases, as well as an algorithm for efficient fence insertion.

To establish our reduction, we first provide a complete characterization of non-robustness in terms of so-called *feasible attacks*. An attack is a pair of load and store instructions of a thread, called the attacker, whose reordering can lead to a non-SC computation. In that case we say the attack is feasible, because it has a (TSO) witness computation. The special form of witness computations then allows us to detect them by tracking SC computations of an *instrumented program*. Given a potential attack, we show how to check its feasibility by solving an SC state reachability query in a linear-size instrumented program. The fact that *only SC semantics* (of the instrumented program) needs to be considered for detecting non-SC behaviors (of the original program) is important: it relieves us from examining TSO computations, which obliges to encode (somehow) the contents of store buffers (as in, e.g., [9,10]). Interestingly, the detection of feasible attacks can be parallelized, which speeds up the decision procedure. Overall, we provide a general reduction of the TSO robustness problem to a quadratic number (in the size of the program) of state reachability queries under the SC semantics in linear-size instrumented programs of the same type as the original one. Our construction is source-to-source and is valid for (1) an arbitrary number of memory addresses/variables, (2) an arbitrary data domain, (3) an arbitrary number of threads, and (4) unbounded store buffers.

With this reduction, we can harness all available techniques and tools for solving reachability queries (either exactly, or approximately) in various classes of concurrent programs, regardless of decidability and complexity issues. It also yields decision algorithms for significant classes of programs. Assume we have a finite number of memory addresses, and the data domain is finite. Then, for a fixed number of threads, a direct consequence of our reduction is that the robustness problem is decidable and in PSPACE since it is polynomially reducible to state reachability in finite-state concurrent programs [18]. Therefore, we obtain in this case a *simple* robustness checking algorithm which matches the complexity upper bound proved in [8]. Our construction also provides an effective decision algorithm for the up to now open case where the *number of threads is arbitrarily large*. Indeed, state reachability queries in this case can be solved in vector addition systems with states (VASS), or equivalently as coverability in Petri nets, which is known to be decidable [26] and EXPSpace-hard [23]. In both cases (fixed or arbitrary number of threads) the decision algorithms do not assume bounded store buffers.

As last contribution, we address the issue of enforcing robustness by fence insertion. Obviously, inserting a fence after each store ensures robustness, but it also ruins all performance benefits that a relaxed memory model brings. A natural requirement on the set of fences is irreducibility, i.e., minimality wrt. set inclusion. Since there may be several irreducible sets enforcing robustness, it is natural to ask for a set that optimizes some notion of cost. We assume that we have a *cost function* that defines the cost of inserting a fence at each program

location. For instance, by assuming cost 1 for all locations, we optimize the size of the fence set. Other cost functions reflect the performance of the resulting program. We propose an algorithm which, given a cost function, computes an optimal set of fences. The algorithm is based on 0/1-integer linear programming and exploits the notion of attacks to guide the selection of fences.

We implemented the algorithms (using SPIN as a back-end reachability checker), and applied them successfully to several examples, including mutual exclusion protocols and concurrent data structures. The experiments we have carried out show that our approach is quite effective: (1) Many of the attacks to be examined can be discarded by simple syntactic checks (e.g., the presence of a fence between the store and load instructions), and those that require solving reachability queries are handled in few seconds, (2) the fence insertion procedure finds efficiently optimal sets of fences, in particular, it can handle the version of the Non-Blocking Write protocol [17] described in [24] (where the write is guarded by a Linux x86 spinlock) for which Owens' method based on so-called *triangular data races* (see related work below) inserts unnecessary fences.

Related work: There are only few results on decidability and complexity of relaxed memory models. Reachability under TSO has been shown to be decidable but non-primitive recursive [7] in the case of a finite number of threads and a finite data domain. In the same case, trace-robustness has been shown to be PSPACE-complete in [8] using a combinatorial approach. The method we adopt in this paper is conceptually and technically different from the one we took in [8]. While we reuse from [8] the fact that it is possible to reason on TSO computations where only one thread has reordered its actions, we develop here a new approach where the main technical contributions reside in the characterization of non-robustness in terms of existence of feasible attacks and in the instrumentation we provide to reduce trace-robustness to SC state reachability. Besides getting decidability and complexity results, this reduction allows to leverage all the existing verification methods and tools for checking (SC) state reachability in various classes of programs to tackle the issue of checking and enforcing robustness against TSO.

Alur et al. have shown that checking sequential consistency of a concurrent implementation wrt. a specification is undecidable in general [6]. This result does not contradict our findings: we consider here the special case where the implementation is the TSO semantics and the specification is the SC semantics of a program. In [14], the problem of deciding whether a given computation is SC feasible has been proved NP-complete. Robustness is concerned with all TSO computations, instead.

As mentioned above, the problem of checking and enforcing trace-robustness against weak memory models has been addressed in [9,10,5], but none of these works provide (sound and complete) decision procedures. Owens proposes in [24] a notion of robustness that is *stronger* than trace-robustness, based on detecting triangular data races. This allows for sound trace-robustness checking but, as

Owens shows in his paper, in some cases leads to unnecessary fences which can be harmful for performance. Moreover, the notion of triangular data races defined in [24] comes without an algorithm for checking it¹. Complexity considerations (using the techniques in [8]) show that detecting triangular data races requires solving state reachability queries under SC. Therefore, as we show here, checking trace-robustness is not more expensive than detecting triangular data races.

State-based robustness (which preserves the reachable states) is a weaker robustness criterion, but does not preserve path properties in contrast to trace-robustness, and is of significantly higher complexity (non-primitive recursive as it can be deduced from [7], instead of PSPACE). It has been addressed in a precise manner in [2] where a symbolic decision procedure together with a fence insertion algorithm are provided. The same issue is addressed in [19,20] using over-approximate reachability analysis based on abstractions of the store buffers.

Finally, let us mention work that considers the dual approach: starting from a robust program, remove unnecessary fences [29]. This work is aimed at compiler optimisations and does not provide a decision procedure for robustness. It can also not find an optimal set of fences the enforce trace-robustness.

2 Parallel Programs

Syntax We consider parallel programs with shared memory as defined by the grammar in Figure 1. Programs have a name and consist of a finite number of threads. Each thread has an identifier and a list of local registers it operates on. The thread’s source code is given as a finite sequence of labelled instructions. More than one instruction can be marked by the same label; this allows us to mimic expressive constructs like conditional branching and iteration with a lightweight syntax. The instruction set includes loads from memory to a local register, stores to memory, memory fences to control TSO store buffers, local computations, and assertions. Figure 2 provides a sample program.

We assume a program comes with two sets: a *data domain* DOM and a *function domain* FUN. The data domain should contain value zero: $0 \in \text{DOM}$. Moreover, we assume that all values from DOM can be used as addresses. Each memory location accessed by loads and stores is identified by such an address, and memory locations identified by different addresses do not overlap. The set FUN contains functions that are defined on the data domain and can be used in the program. Note that we do not make any finiteness assumptions.

TSO Semantics Fix a program \mathcal{P} with threads $\text{THRD} := \{t_1, \dots, t_n\}$. Let each thread t_i have the initial label $l_{0,i}$ and declare registers \bar{r}_i . We define the set of variables as the union of addresses and registers: $\text{VAR} := \text{DOM} \cup \cup_{i \in [1,n]} \bar{r}_i$. We denote the set of all instruction labels that occur in threads by LAB.

The TSO semantics is operational, in terms of states and labelled transitions between them. On the x86 TSO architecture, each processor effectively has a local FIFO buffer that keeps stores for later execution [25,27,9,10]. Therefore,

¹ Citation from [24]: “... formal reasoning directly on traces can be tedious, so a program logic or sound static analyzer specialized to proving triangular-race freedom might make the application of TRF more convenient.”

```

⟨prog⟩ ::= program ⟨pid⟩ ⟨thrd⟩*
⟨thrd⟩ ::= thread ⟨tid⟩
          regs ⟨reg⟩*
          init ⟨label⟩
          begin ⟨linst⟩* end
⟨linst⟩ ::= ⟨label⟩: ⟨inst⟩; goto ⟨label⟩;
⟨inst⟩ ::= ⟨reg⟩ ← mem[⟨expr⟩]
          | mem[⟨expr⟩] ← ⟨expr⟩
          | mfence
          | ⟨reg⟩ ← ⟨expr⟩
          | assert ⟨expr⟩
⟨expr⟩ ::= ⟨fun⟩(⟨reg⟩*)

```

Fig. 1: Syntax of parallel programs.

```

program Dekker
thread t1 regs r1 init l0 begin
l0: mem[x] ← 1; goto l1;
l1: r1 ← mem[y]; goto l2;
end
thread t2 regs r2 init l'0 begin
l'0: mem[y] ← 1; goto l'1;
l'1: r2 ← mem[x]; goto l'2;
end

```

Fig. 2: Simplified version of Dekker's mutex algorithm. Under SC, it is impossible that $r_1 = r_2 = 0$ when both threads reach l_2 and l'_2 .

a *state* is a triple $s = (\text{pc}, \text{val}, \text{buf})$ where program counter $\text{pc}: \text{THRD} \rightarrow \text{LAB}$ holds, for each thread, the label of the instruction(s) to be executed next. The valuation $\text{val}: \text{VAR} \rightarrow \text{DOM}$ gives the values of registers and memory locations. The third component $\text{buf}: \text{THRD} \rightarrow (\text{DOM} \times \text{DOM})^*$ is the (per thread) buffer content: a sequence of address-value pairs $a \leftarrow v$.

In the *initial state* $s_0 := (\text{pc}_0, \text{val}_0, \text{buf}_0)$ the program counter is set to the initial labels, $\text{pc}_0(t_i) = l_{0,i}$ for all $t_i \in \text{THRD}$, registers and addresses hold value zero, $\text{val}_0(x) = 0$ for all $x \in \text{VAR}$, and all buffers are empty, $\text{buf}_0(t) := \varepsilon$ for all $t \in \text{THRD}$. Here, ε denotes the empty sequence.

Instructions yield transitions between states that are labelled by *actions* from $\text{ACT} := \text{THRD} \times (\{\text{isu}, \text{loc}\} \cup (\{\text{ld}, \text{st}\} \times \text{DOM} \times \text{DOM}))$. An action consists of the thread id and the actual arguments of an executed instruction. We use *loc* to abstract assignments and asserts that are local to the thread. The issue action *isu* indicates that a store was executed on the processor. The store action (t, st, a, v) gives the moment when the store becomes visible in memory.

The *TSO transition relation* \rightarrow_{TSO} is the smallest relation between TSO states that is defined by the rules in Table 1. The rules repeat, up to notation and support for locked instructions, Figure 1 from [25]. The first two rules implement loads from the buffer and from the memory respectively. By the third rule, store instructions enqueue write operations to the buffer. The fourth rule non-deterministically dequeues and executes them on memory. The fifth rule defines that memory fences can only be executed when the buffer is empty. The last two rules refer to local assignments and assertions. We omitted locked instructions to keep things simple. Their introduction is straightforward and does not affect the results. Indeed, our implementation supports them [1].

The set of *TSO computations* contains all sequences of actions that lead from the initial TSO state to a state where all buffers are empty:

$$C_{\text{TSO}}(\mathcal{P}) := \{\tau \in \text{ACT}^* \mid s_0 \xrightarrow{\tau}_{\text{TSO}} s \text{ for some TSO state } s = (\text{pc}, \text{val}, \text{buf}) \text{ with } \text{buf}(t) = \varepsilon \text{ for all } t \in \text{THRD}\}.$$

$$\begin{array}{c}
\frac{\langle instr \rangle = r \leftarrow \text{mem}[f_a(\bar{r}_a)], a = f_a(\text{val}(\bar{r}_a)), \text{buf}(t) \downarrow (a \leftarrow *) = \beta \cdot (a \leftarrow v)}{(\text{pc}, \text{val}, \text{buf}) \xrightarrow{(t, \text{ld}, a, v)}_{\text{TSO}} (\text{pc}', \text{val}[r := v], \text{buf})} \\
\frac{\langle instr \rangle = r \leftarrow \text{mem}[f_a(\bar{r}_a)], a = f_a(\text{val}(\bar{r}_a)), \text{buf}(t) \downarrow (a \leftarrow *) = \varepsilon, v = \text{val}(a)}{(\text{pc}, \text{val}, \text{buf}) \xrightarrow{(t, \text{ld}, a, v)}_{\text{TSO}} (\text{pc}', \text{val}[r := v], \text{buf})} \\
\frac{\langle instr \rangle = \text{mem}[f_a(\bar{r}_a)] \leftarrow f_v(\bar{r}_v), a = f_a(\text{val}(\bar{r}_a)), v = f_v(\text{val}(\bar{r}_v))}{(\text{pc}, \text{val}, \text{buf}) \xrightarrow{(t, \text{isu})}_{\text{TSO}} (\text{pc}', \text{val}, \text{buf}[t := \text{buf}(t) \cdot (a \leftarrow v)])} \\
\frac{\text{buf}(t) = (a \leftarrow v) \cdot \beta}{(\text{pc}, \text{val}, \text{buf}) \xrightarrow{(t, \text{st}, a, v)}_{\text{TSO}} (\text{pc}, \text{val}[a := v], \text{buf}[t := \beta])} \\
\frac{\langle instr \rangle = \text{mfence}, \text{buf}(t) = \varepsilon}{(\text{pc}, \text{val}, \text{buf}) \xrightarrow{(t, \text{loc})}_{\text{TSO}} (\text{pc}', \text{val}, \text{buf})} \\
\frac{\langle instr \rangle = r \leftarrow f(\bar{r})}{(\text{pc}, \text{val}, \text{buf}) \xrightarrow{(t, \text{loc})}_{\text{TSO}} (\text{pc}', \text{val}[r := f(\text{val}(\bar{r}))], \text{buf})} \\
\frac{\langle instr \rangle = \text{assert } f(\bar{r}), f(\text{val}(\bar{r})) \neq 0}{(\text{pc}, \text{val}, \text{buf}) \xrightarrow{(t, \text{loc})}_{\text{TSO}} (\text{pc}', \text{val}, \text{buf})}
\end{array}$$

Table 1: TSO transition rules, assuming $\text{pc}(t) = l$, an instruction $\langle instr \rangle$ at label l with destination l' , and $\text{pc}' := \text{pc}[t := l']$. We use \downarrow to denote projection and $*$ for any value, i.e., $\text{buf}(t) \downarrow (a \leftarrow *)$ is a list of address-value pairs in the buffer of thread t having the address a .

The requirement of empty buffers is not important for our results but rather a modelling choice. Figure 3 presents a TSO computation of Dekker's program where the store of the first thread is delayed past the load.

$$\tau = (t_1, \text{isu}) (t_1, \text{ld}, y, 0) (t_2, \text{isu}) (t_2, \text{st}, y, 1) (t_2, \text{ld}, x, 0) (t_1, \text{st}, x, 1)$$

Fig. 3: A TSO computation of Dekker's algorithm. Actions drawn in red belong to the first thread, actions in blue belong to the second thread. The arc connects the issue action with the corresponding delayed store action of the first thread.

SC Semantics Under SC [21], stores are not buffered and hence states are pairs (pc, val) . The rules for SC transitions are appropriately simplified TSO rules. To avoid case distinctions between TSO and SC in the definition of traces, a store instruction generates two actions: an issue followed by the store. Memory fences have no effect under SC. We denote the set of all SC computations of \mathcal{P} by

$$\text{C}_{\text{SC}}(\mathcal{P}) := \{\sigma \in \text{ACT}^* \mid s_0 \xrightarrow{\sigma}_{\text{SC}} s \text{ for some SC state } s\}.$$

3 TSO Robustness

Robustness ensures that the behaviour of a program does not change when it is run on TSO hardware as compared to SC. We study trace-based robustness as in [28,9,10,5,8]. Traces capture the essence of a computation: the control and data dependencies among actions. More formally, consider some computation $\alpha \in C_{SC}(\mathcal{P}) \cup C_{TSO}(\mathcal{P})$. The *trace* $\text{Tr}(\alpha)$ is a graph where the nodes are labelled by the actions in α (stores and issue yield one node). The arcs are defined by the following relations. We have a per thread $t \in \text{THRD}$ total order \rightarrow_{po}^t that gives the order in which the actions of t where issued. Similarly, we have a per address $a \in \text{DOM}$ total order \rightarrow_{st}^a that gives the ordering of all stores to a . We call the unions $\rightarrow_{po} := \cup_{t \in \text{THRD}} \rightarrow_{po}^t$ and $\rightarrow_{st} := \cup_{a \in \text{DOM}} \rightarrow_{st}^a$ the *program order* and the *store order* of the trace. Finally, there is a *source relation* \rightarrow_{src} that determines the store from which a load receives its value. By $\text{Tr}_{mm}(\mathcal{P}) := \text{Tr}(C_{mm}(\mathcal{P}))$ with $mm \in \{SC, TSO\}$ we denote the set of all *SC/TSO traces* of program \mathcal{P} . The *TSO robustness problem* checks whether the sets coincide.

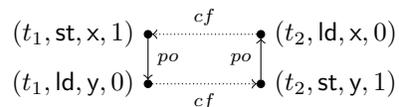
Given: A parallel program \mathcal{P} .

Problem: Does $\text{Tr}_{TSO}(\mathcal{P}) = \text{Tr}_{SC}(\mathcal{P})$ hold?

Since inclusion $\text{Tr}_{SC}(\mathcal{P}) \subseteq \text{Tr}_{TSO}(\mathcal{P})$ always holds, we only have to check the reverse inclusion. We call a computation $\tau \in C_{TSO}(\mathcal{P})$ *violating* if its trace is not among the SC traces of the program, i.e., $\text{Tr}(\tau) \notin \text{Tr}_{SC}(\mathcal{P})$. Violating TSO-computations employ cyclic accesses to addresses that SC is unable to serialize [28]. The cyclic accesses are made visible using a *conflict relation* from loads to stores. Intuitively, $\text{ld} \rightarrow_{cf} \text{st}$ means that st overwrites a value that ld reads. The union of all four relations is commonly called *happens-before relation* of the trace, $\rightarrow_{hb} := \rightarrow_{po} \cup \rightarrow_{st} \cup \rightarrow_{src} \cup \rightarrow_{cf}$.

Lemma 1 ([28]). *Consider TSO trace $\text{Tr}(\tau) \in \text{Tr}_{TSO}(\mathcal{P})$. Then $\text{Tr}(\tau) \in \text{Tr}_{SC}(\mathcal{P})$ iff \rightarrow_{hb} is acyclic.*

Consider the computation in Figure 3. The load from thread t_1 conflicts with the store from t_2 because the load reads the initial value of y while the store overwrites it. The situation with the load from t_2 and the store from t_1 is symmetric. Together with the program order, the conflict relations produce a cycle:



Indeed, there is no SC computation with this trace, as predicted by Lemma 1.

Lemma 1 does not provide a method for finding cyclic traces. We have recently shown that TSO robustness is decidable, in fact, PSPACE-complete [8]. The algorithm underlying this result, however, is based on enumeration and not meant to be implemented. The main contribution of the present work is a novel and practical approach to checking robustness.

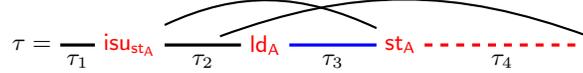


Fig. 4: TSO witness for the attack $(t_A, \text{stinst}, \text{ldinst})$. It satisfies the following constraints. **(W1)** Only the attacker delays stores. **(W2)** Store st_A is an instance of stinst . It is the first store of the attacker that is delayed. Load ld_A is an instance of ldinst . It is the last action of the attacker that is overstepped by st_A . So τ_2 contains loads, assignments, asserts, and issues, but no fences and stores of the attacker. It may contain arbitrary helper actions. **(W3)** We require $\text{ld}_A \rightarrow_{\text{hb}}^+ \text{act}$ for every action act in $\text{ld}_A \cdot \tau_3 \cdot \text{st}_A$. An issue + store of a helper is counted as one action act . **(W4)** Sequence τ_4 only consists of stores of the attacker that were issued before ld_A and that have been delayed. **(W5)** All these stores st satisfy $\text{addr}(\text{st}) \neq \text{addr}(\text{ld}_A)$, i.e., ld_A has not read its value early.

The only concept we keep from our earlier work are minimal violations. A *minimal violation* is a violating computation that uses a minimal total number of delays. Interestingly, for minimal violations the following holds.

Lemma 2 (Locality [8], Appendix B). *In a minimal violation, only a single thread delays stores.*

Consider the computation in Figure 3. It relies on a single delay in thread t_1 and, indeed, is a minimal violation. As predicted by the lemma, the second thread writes to its buffer and immediately flushes it.

4 Attacks on TSO Robustness

Our approach to checking TSO robustness combines two insights. We first rephrase robustness in terms of a simpler problem: the absence of feasible attacks. We then devise an algorithm that checks attacks for feasibility. Interestingly, SC reachability techniques are sufficient for this purpose. Together, this yields a sound and complete reduction of TSO robustness to SC reachability.

The notion of attacks is inspired by the shape of minimal violations. We show that if a program is not robust, then there are violations of the form shown in Figure 4: one thread, the *attacker*, delays a store action st_A past a later load action ld_A in order to break robustness. The remaining threads become *helpers* and provide a happens-before path from ld_A to st_A . This yields a happens-before cycle and shows non-robustness.

Thread, store instruction stinst of st_A , and load instruction ldinst of ld_A are syntactic objects. The idea of our approach is to fix these three parameters, the *attack*, prior to the analysis. The algorithm then tries to find a witness computation that proves the attack feasible.

Definition 1. *An attack $A = (t_A, \text{stinst}, \text{ldinst})$ consists of a thread $t_A \in \text{THRD}$ called attacker, a store instruction stinst and a load instruction ldinst . A TSO witness for A is a computation of the form in Figure 4, i.e., it satisfies **(W1)** to **(W5)**. If a TSO witness exists, the attack is called feasible.*

In Dekker’s algorithm, there is an attack $A = (t_A, \text{stinst}, \text{ldinst})$ with $t_A = t_1$, stinst the store at l_0 , and ldinst the load at l_1 . A TSO witness of this attack is

the computation τ from Figure 3. With reference to Figure 4 we have $\tau_1 = \varepsilon$, $\text{isu}_{\text{st}_A} = (t_1, \text{isu})$, $\tau_2 = \varepsilon$, $\text{ld}_A = (t_1, \text{ld}, y, 0)$, $\tau_3 = (t_2, \text{isu}) \cdot (t_2, \text{st}, y, 1) \cdot (t_2, \text{ld}, x, 0)$, $\text{st}_A = (t_1, \text{st}, x, 1)$, $\tau_4 = \varepsilon$. The program also contains a symmetric attack A' with t_2 as the attacker.

Although TSO witnesses are quite restrictive computations, robustness can be reduced to verifying that no attack has a TSO witness.

Theorem 1 (Complete Characterization of Robustness with Attacks).
Program \mathcal{P} is robust iff no attack is feasible, i.e., no attack admits a TSO witness.

Proof. The existence of a TSO witness implies non-robustness of the program. Indeed, a TSO witness comes with a happens-before cycle $\text{st}_A \xrightarrow{+}_{\text{po}} \text{ld}_A \xrightarrow{+}_{\text{hb}} \text{st}_A$. We argue that also the reverse holds: if a program is not robust, there is a feasible attack. Assume \mathcal{P} is not robust. We construct a TSO witness computation. Among the violating computations, we select $\tau \in C_{\text{TSO}}(\mathcal{P})$ where the number of delays is minimal. The computation need not be unique. By Lemma 2, in τ only one thread t_A uses its buffer and **(W1)** holds. We elaborate on the shape of τ .

Initially, the attacker executes under SC so that stores immediately follow their issues. This computation is embedded into τ_1 in Figure 4. Eventually, the attacker starts delaying stores. Let st_A be the first store that is delayed. It gets reordered past several loads, the last of which being ld_A . This shows **(W2)**.

The helper actions in τ_3 are depicted in blue in Figure 4. To see that we can assume **(W3)**, first note that $\text{ld}_A \xrightarrow{+}_{\text{hb}} \text{st}_A$ holds. If there was no such path, st_A could be placed before ld_A without changing the trace. This would save a delay, in contradiction to minimality of τ . Assume $\tau_3 = \tau'_3 \cdot \text{act} \cdot \tau''_3$ where act is the first action so that $\text{ld}_A \not\xrightarrow{+}_{\text{hb}} \text{act}$. Then act is independent from all actions in $\text{ld}_A \cdot \tau'_3$. We find a computation with the same trace where act is placed before ld_A .

With cycle $\text{st}_A \xrightarrow{+}_{\text{po}} \text{ld}_A \xrightarrow{+}_{\text{hb}} \text{st}_A$, computation τ_4 only needs to contain the stores of the attacker that have been delayed past ld_A . Since these stores are non-blocking, the helpers can stop with the last action in τ_3 . We can moreover assume ld_A to be the program order last action of the attacker. **(W4)** holds.

We now argue that ld_A has not read its value early from any of the delayed stores, **(W5)**. Towards a contradiction, assume ld_A obtained its value from st in $\tau_4 = \tau_{41} \cdot \text{st} \cdot \tau_{42}$. There is a computation τ' where we avoid the early read: it replaces τ_4 by $\tau_{41} \cdot \text{st} \cdot \text{ld}_A \cdot \tau_{42}$. The traces of τ and τ' coincide, but τ' saves the delay of st past ld_A . A contradiction to minimality.

It is readily checked that τ is a TSO witness for the attack $(t_A, \text{stinst}, \text{ldinst})$ where stinst and ldinst are the instructions that st_A and ld_A are derived from. \square

Since the number of attacks is only quadratic in the size of the program, we can just enumerate them and check whether one admits a TSO witness. To check whether a witness exists, we employ the instrumentation described in the following section.

5 Instrumentation

Consider program \mathcal{P} with attack $A = (t_A, \text{stinst}, \text{ldinst})$. TSO witnesses for A only make limited use of buffers, to an extent that allows us to characterize them

by SC computations in a program \mathcal{P}_A that is *instrumented for attack A*. By instrumentation we mean that \mathcal{P}_A replaces every thread by a modified version. Capturing TSO witnesses with a program that runs under SC is difficult for two reasons. First, TSO has unbounded store buffers which can delay stores arbitrarily long. Second, the happens-before dependence that the helpers create may involve an arbitrary number of actions. Our instrumentation copes with both problems using the following tricks.

To handle store buffering, we instrument the attacker thread (Section 5.1). Essentially, we emulate store buffering under SC using auxiliary addresses. To explain the idea, consider the TSO witness in Figure 4. When the instrumented attacker executes the delayed stores $st_A \cdot \tau_4$ under SC, they occur right behind their issue actions. To mimic store buffering, these stores now access auxiliary addresses that the other threads do not load. As a result, the stores remain invisible to the helpers. This is as intended: the delayed stores $st_A \cdot \tau_4$ in Figure 4 are also never accessed by helper threads. But how many auxiliary addresses do we need to faithfully simulate buffers? It is sufficient to have *a single auxiliary address* per address in the program. The reason is that a load always reads the most recent store to its address that is held in the buffer.

To build up a happens-before path from ld_A to st_A , we instrument the helper threads (Section 5.2). The question is how to decide whether a new action act is in happens-before relation with an earlier action act' so that $ld_A \rightarrow_{hb}^* act' \rightarrow_{hb}^* act$. What is the information we need about the earlier actions in order to append act ? It is sufficient to know two facts. Has the thread already contributed an action act' ? This information ensures $act' \rightarrow_{po}^* act$, and can be kept in the control flow of the thread. Moreover, we keep track of whether the path contains a load or store access to the address $addr(act)$. If there was a load access $act' = ld$, we can add a store $act = st$ and get $ld \rightarrow_{hb}^* st$. If there was a store, we are free to add a load or a store. Hence, we need *one auxiliary address* per address in the program for this access information: no access, load access, store access.

Consider the TSO witness for Dekker given in Figure 3. Instead of buffering $(t_1, st, x, 1)$, the instrumentation immediately executes the store after its issue action. But instead of address x , the action accesses the auxiliary address (x, d) that the other threads do not load. To indicate that this store is invisible to the helper threads, we depict it in gray. So, the SC computation of the instrumented program roughly looks like this:

$$(t_1, isu) \cdot (t_1, st, (x, d), 1) \cdot (t_1, ld, y, 0) \stackrel{(1)}{\cdot} (t_2, isu)(t_2, st, y, 1) \stackrel{(2)}{\cdot} (t_2, ld, x, 0).$$

At moment (1), we know that there has been a load access to address y . At moment (2), address y has even seen a store. At the end of the computation, address y has seen a store and address x has seen a load.

The store of t_2 can be appended since it is in happens-before relation with the attacker's load. The following load can be added as t_2 has contributed the previous store. The search terminates here since the helper's load accesses address x that was used by the store from the attack.

5.1 Instrumentation of the Attacker

The instrumentation emulates the buffering of stores in a TSO witness (Figure 4). Starting from st_A , the stores are replaced by stores st_A^{aux} to auxiliary addresses (a, d) that are only visible to the attacker. As long as a has not been written, (a, d) holds the initial value 0. Once the attacker stores v into a , we set $\text{mem}[(a, d)] = (v, d)$. In this way, (a, d) always holds the most recent store. A load $r \leftarrow \text{mem}[a]$ of the attacker reads a value v from the buffer whenever $\text{mem}[(a, d)] = (v, d)$; otherwise $\text{mem}[(a, d)] = 0$ and the load obtains the value $v = \text{mem}[a]$ from memory. We turn to the translation.

Let t_A declare registers r^* , have initial location l_0 , and define instructions $\langle \text{linst} \rangle^*$ that contain stinst and ldinst from the attack. The instrumentation is

$$\begin{aligned} \llbracket t_A \rrbracket &:= \text{thread } \tilde{t}_A \text{ regs } r^* \text{ init } l_0 \\ &\quad \text{begin } \langle \text{linst} \rangle^* \llbracket \text{stinst} \rrbracket_{A1} \llbracket \text{ldinst} \rrbracket_{A1} \llbracket \langle \text{linst} \rangle^* \rrbracket_{A2} \text{ end.} \end{aligned}$$

It introduces a copy of the source code $\llbracket \langle \text{linst} \rangle^* \rrbracket_{A2}$ where the stores are replaced by accesses to auxiliary addresses. To move to the code copy, the attacker uses an instrumented version $\llbracket \text{stinst} \rrbracket_{A1}$ of stinst .

$$\begin{aligned} \llbracket l_1: \text{mem}[e_1] \leftarrow e_2; \text{goto } l_2; \rrbracket_{A1} &:= l_1: \text{mem}[(e_1, d)] \leftarrow (e_2, d); \text{goto } \tilde{l}_x; & (1) \\ &\quad \tilde{l}_x: \text{mem}[a_{\text{st}_A}] \leftarrow e_1; \text{goto } \tilde{l}_2; \end{aligned}$$

$$\begin{aligned} \llbracket l_1: r \leftarrow \text{mem}[e]; \text{goto } l_2; \rrbracket_{A1} &:= \tilde{l}_1: \text{assert mem}[(e, d)] = 0; \text{goto } \tilde{l}_{x1}; & (2) \\ &\quad \tilde{l}_{x1}: \text{mem}[\text{hb}] \leftarrow \text{true}; \text{goto } \tilde{l}_{x2}; \\ &\quad \tilde{l}_{x2}: \text{mem}[(e, \text{hb})] \leftarrow \text{lda}; \text{goto } \tilde{l}_{x3}; \end{aligned}$$

$$\llbracket l_1: \text{mem}[e_1] \leftarrow e_2; \text{goto } l_2; \rrbracket_{A2} := \tilde{l}_1: \text{mem}[(e_1, d)] \leftarrow (e_2, d); \text{goto } \tilde{l}_2; \quad (3)$$

$$\begin{aligned} \llbracket l_1: r \leftarrow \text{mem}[e]; \text{goto } l_2; \rrbracket_{A2} &:= \tilde{l}_1: \text{assert mem}[(e, d)] = 0; \text{goto } \tilde{l}_{x1}; & (4) \\ &\quad \tilde{l}_{x1}: r \leftarrow \text{mem}[e]; \text{goto } \tilde{l}_2; \\ &\quad \tilde{l}_1: \text{assert mem}[(e, d)] \neq 0; \text{goto } \tilde{l}_{x2}; \\ &\quad \tilde{l}_{x2}: (r, d) \leftarrow \text{mem}[(e, d)]; \text{goto } \tilde{l}_2; \end{aligned}$$

$$\llbracket l_1: \text{local}; \text{goto } l_2; \rrbracket_{A2} := \tilde{l}_1: \text{local}; \text{goto } \tilde{l}_2; \quad (5)$$

$$\llbracket l_1: \text{mfence}; \text{goto } l_2; \rrbracket_{A2} := \quad (6)$$

Fig. 5: Instrumentation of the attacker.

The translation of instructions is defined in Figure 5. We make a few remarks. The instrumentation of $\text{stinst} = l_1: \text{mem}[e_1] \leftarrow e_2; \text{goto } l_2;$ keeps the address used in the store in a fresh address a_{st_A} . For the sake of readability, in Equation (4) we use memory accesses in instructions other than load and store. Equation (6) deletes fences, as they forbid to delay st_A over ld_A . Let $\text{ldinst} = l_1: r \leftarrow \text{mem}[e]; \text{goto } l_2;$ be the load used in the attack. Equation (2) checks that the load has not read its value early and sets an auxiliary happens-before address (e, hb) to access level load, lda . We postpone the definition of

```

thread  $\tilde{l}_1$  regs  $r_1$  init  $l_0$  begin
/* Original code */
 $l_0$ : mem[x]  $\leftarrow$  1; goto  $l_1$ ;
 $l_1$ :  $r_1 \leftarrow$  mem[y]; goto  $l_2$ ;

/* Instrumented stinst */
 $l_0$ : mem[(x,d)]  $\leftarrow$  (1,d); goto  $\tilde{l}_x$ ;
 $\tilde{l}_x$ : mem[ $a_{st_A}$ ]  $\leftarrow$  x; goto  $\tilde{l}_1$ ;

/* Instrumented ldinst */
 $\tilde{l}_1$ : assert mem[(y,d)] = 0; goto  $\tilde{l}_{x1}$ ;
 $\tilde{l}_{x1}$ : mem[hb]  $\leftarrow$  true; goto  $\tilde{l}_{x2}$ ;
 $\tilde{l}_{x2}$ : mem[(y,hb)]  $\leftarrow$  lda; goto  $\tilde{l}_{x3}$ ;
end
/* Instrumented copy of the store */
 $\tilde{l}_0$ : mem[(x,d)]  $\leftarrow$  (1,d); goto  $\tilde{l}_1$ ;

/* Instrumented copy of the load */
 $\tilde{l}_1$ : assert mem[(y,d)] = 0; goto  $\tilde{l}_{x4}$ ;
 $\tilde{l}_{x4}$ :  $r \leftarrow$  mem[y]; goto  $\tilde{l}_2$ ;
 $\tilde{l}_1$ : assert mem[(y,d)]  $\neq$  0; goto  $\tilde{l}_{x5}$ ;
 $\tilde{l}_{x5}$ : (r,d)  $\leftarrow$  mem[(y,d)]; goto  $\tilde{l}_2$ ;

```

Fig. 6: Attacker instrumentation of thread t_1 in Dekker from Figure 2. The attacker's store is the store at label l_0 , the load is the load at label l_1 .

access levels until the translation of helpers. It also sets `hb` flag for helpers to indicate that they cannot execute actions not contributing to the happens-before path. Figure 6 illustrates the instrumentation on our running example.

5.2 Instrumentation of Helpers

In TSO witnesses, by **(W3)**, all helper actions after ld_A are in happens-before relation with ld_A . To ensure this, we make use of Lemma 3. The proof from left to right is by definition of happens before. For the reverse direction, note that happens-before is *stable under insertion*. Consider $st \rightarrow_{src} ld$. A happens-before relation remains valid in any computation that places actions between st and ld .

Lemma 3. Consider $\tau = \tau_1 \cdot act_1 \cdot \tau_2 \in C_{SC}(\mathcal{P})$ where for all act_2 in τ_2 we have $act_1 \rightarrow_{hb}^* act_2$. Computation $\tau \cdot act$ satisfies $act_1 \rightarrow_{hb}^* act$ iff

- (i) there is an action act_2 in $act_1 \cdot \tau_2$ with $thread(act_2) = thread(act)$ or
- (ii) act is a load whose address is stored in $act_1 \cdot \tau_2$ or
- (iii) act is a store (with issue) whose address is loaded or stored in $act_1 \cdot \tau_2$.

The lemma suggests the following instrumentation. For every helper t , we track whether it has executed an action that depends on ld_A . The idea is to use the control flow. Upon detection of this first action, the thread moves to a copy of its code. All actions from this copy stay in happens-before relation with ld_A .

It remains to decide whether an action act allows a thread to move to the code copy. According to Lemma 3, this depends on the earlier accesses to the address $a = addr(act)$. We introduce auxiliary *happens-before addresses* (a, hb) that provide this access information. The addresses (a, hb) range over the domain $\{0, lda, sta\}$ of *access types*. It is sufficient to keep track of the *maximal* access type wrt. the ordering 0 (no access) $<$ lda (load access) $<$ sta (store access).

$$\llbracket l_1 : instr ; goto l_2 ; \rrbracket_{H0} := l_1 : \text{assert mem}[\text{hb}] = 0 ; goto l_x ; \quad (7)$$

$$l_x : instr ; goto l_2 ;$$

$$\llbracket l_1 : r \leftarrow \text{mem}[e] ; goto l_2 ; \rrbracket_{H1} := l_1 : \text{assert mem}[(e, \text{hb})] = \text{sta} ; goto \tilde{l}_x ; \quad (8)$$

$$\tilde{l}_x : r \leftarrow \text{mem}[e] ; goto \tilde{l}_2 ;$$

$$\llbracket l_1 : \text{mem}[e_1] \leftarrow e_2 ; goto l_2 ; \rrbracket_{H1} := l_1 : \text{assert mem}[(e_1, \text{hb})] \geq \text{lda} ; goto \tilde{l}_{x1} ; \quad (9)$$

$$\tilde{l}_{x1} : \text{mem}[e_1] \leftarrow e_2 ; goto \tilde{l}_{x2} ;$$

$$\tilde{l}_{x2} : \text{mem}[(e_1, \text{hb})] \leftarrow \text{sta} ; goto \tilde{l}_2 ;$$

$$\llbracket l_1 : local/mfence ; goto l_2 ; \rrbracket_{H2} := \tilde{l}_1 : local/mfence ; goto \tilde{l}_2 ; \quad (10)$$

$$\llbracket l_1 : \text{mem}[e_1] \leftarrow e_2 ; goto l_2 ; \rrbracket_{H2} := \tilde{l}_1 : \text{mem}[e_1] \leftarrow e_2 ; goto \tilde{l}_e ; \quad (11)$$

$$\tilde{l}_e : \text{mem}[(e_1, \text{hb})] \leftarrow \text{sta} ; goto \tilde{l}_2 ;$$

$$\llbracket l_1 : r \leftarrow \text{mem}[e] ; goto l_2 ; \rrbracket_{H2} := \tilde{l}_1 : \tilde{r} \leftarrow e ; goto \tilde{l}_{x1} ; \quad (12)$$

$$\tilde{l}_{x1} : r \leftarrow \text{mem}[\tilde{r}] ; goto \tilde{l}_{x2} ;$$

$$\tilde{l}_{x2} : \text{mem}[(\tilde{r}, \text{hb})] \leftarrow \max\{\text{lda}, \text{mem}[(\tilde{r}, \text{hb})]\} ; goto \tilde{l}_2 ;$$

$$\llbracket \langle l \rangle \rrbracket_{H3} := \tilde{l} : \tilde{r} \leftarrow \text{mem}[a_{\text{st}_A}] ; goto \tilde{l}_{x1} ; \quad (13)$$

$$\tilde{l}_{x1} : \tilde{r} \leftarrow \text{mem}[(\tilde{r}, \text{hb})] ; goto \tilde{l}_{x2} ;$$

$$\tilde{l}_{x2} : \text{assert } \tilde{r} \neq 0 ; goto \tilde{l}_{x3} ;$$

$$\tilde{l}_{x3} : \text{mem}[\text{suc}] \leftarrow \text{true} ; goto \tilde{l}_{x4} ;$$

Fig. 7: Instrumentation of helpers.

For the definition, consider a helper thread t that declares r^* , has initial label l_0 , and defines instructions $\langle linst \rangle^*$. The instrumented thread is

$$\begin{aligned} \llbracket t \rrbracket &:= \text{thread } \tilde{t} \text{ regs } \tilde{r}, r^* \text{ init } l_0 \\ &\quad \text{begin } \llbracket \langle linst \rangle \rrbracket_{H0} * \llbracket \langle ldinst \rangle \rrbracket_{H1}^* \llbracket \langle linst \rangle \rrbracket_{H2}^* \llbracket \langle l \rangle \rrbracket_{H3}^* \text{ end.} \end{aligned}$$

Here, $\langle ldinst \rangle^*$ is the subsequence of all load and store instructions. Their instrumentation $\llbracket \langle ldinst \rangle \rrbracket_{H1}^*$ is used to move to the code copy $\llbracket \langle linst \rangle \rrbracket_{H2}^*$. Moreover, $\langle l \rangle^*$ are all labels used by the thread. The additional instructions $\llbracket \langle l \rangle \rrbracket_{H3}^*$ raise a success flag when a TSO witness has been found. $\llbracket \langle linst \rangle \rrbracket_{H0}$ forces helpers to either enter the code copy or stop when hb flag is raised.

The translation of instructions is given in Figure 7. We make some remarks. Transitions to the code copy check the auxiliary addresses for whether the current action is in happens-before relation with ld_A . Loads in Equation (8) check for an earlier store access to their address, Lemma 3(ii). Stores in Equation (9) require that the address has seen at least a load, Lemma 3(iii). They set the access level to sta . Loads and stores in the code copy maintain the auxiliary addresses to contain the maximal access types, Equations (12) and (11). Note the auxiliary register \tilde{r} that ensures we do not overwrite the address. At every label of the code copy we add a check, Equation (13), whether the address used in the attack's store has been accessed in the code copy. If so, a success flag is raised.

5.3 Soundness and Completeness

The flag indicates that the SC computation corresponds to a TSO witness, and we call (pc, val) with $val(suc) = true$ a *goal configuration*. The instrumentation thus reduces feasibility of attack A to SC reachability of a goal configuration in program \mathcal{P}_A . The instrumentation is sound and complete. If a goal configuration is reachable, we can reconstruct a TSO witness for the attack. In turn, every TSO witness ensures the goal configuration is reachable.

Theorem 2 (Soundness and Completeness). *Attack $A = (t_A, stinst, ldinst)$ is feasible in program \mathcal{P} iff program \mathcal{P}_A reaches a goal configuration under SC.*

In combination with Theorem 1, we can check robustness by inspecting all \mathcal{P}_A .

Theorem 3 (From Robustness to SC Reachability). *Program \mathcal{P} is robust iff no instrumentation \mathcal{P}_A reaches a goal configuration under SC.*

The instrumentation we provide is linear in size. Then, it follows from Theorem 3 that checking robustness for programs over finite data domains is in PSPACE. The problem is actually PSPACE-complete due to the lower bound in [8].

6 Robustness for Parameterized Programs

We extend the study of robustness to *parameterized programs*. A parameterized program represents an infinite family of instance programs that replicate the threads multiple times. Syntactically, parameterized programs coincide with the parallel programs we introduced in Section 2: they have a name and declare a finite set of threads t_1, \dots, t_k . The difference is in the semantics. A parameterized program represents a family of programs: for every vector $I = (n_1, \dots, n_k) \in \mathbb{N}^k$, a *program instance* $\mathcal{P}(I)$ declares n_i copies of thread t_i .

In the parameterized setting, the robustness problem asks whether all instances of a given program are robust:

Given: A parameterized program \mathcal{P} .

Problem: Does $\text{Tr}_{\text{TSO}}(\mathcal{P}(I)) = \text{Tr}_{\text{SC}}(\mathcal{P}(I))$ hold for all instances $\mathcal{P}(I)$ of \mathcal{P} ?

The problem is interesting because libraries usually cannot make assumptions on the number of threads that use their functions. They have to guarantee proper functioning for any number.

We reduce robustness for parameterized programs to a parameterized version of reachability, based on the following insight. A parameterized program is not robust if and only if there is an instance $\mathcal{P}(I)$ that is not robust. With Theorem 1, instance $\mathcal{P}(I)$ is not robust if and only if there is an attack A that is feasible. With the instrumentation from Section 5 and Theorem 3, this feasibility can be checked as reachability of a goal configuration in $\mathcal{P}(I)_A$.

Algorithmically, it is impossible to instrument all (infinitely many) instance programs. Instead, the idea is to instrument directly the parameterized program \mathcal{P} towards the attack A . Using the constructions from Section 5, we modify every thread and again obtain program \mathcal{P}_A , which is now parameterized.

Actually, for the attacker we have to be slightly more careful. In an instance program, only one copy of the thread should act as attacker, the remaining copies have to behave like helpers. Therefore, the thread must be instrumented not only as an attacker, but also as a helper. To ensure that only one copy of the attacker delays stores, we add an additional flag variable. Before starting an attack, the thread checks this variable. If it contains the initial value, the thread sets the flag and starts delaying stores. If it has a different value, the thread continues to run sequentially. This check requires an atomic test-and-set operation which can be implemented on x86 by the `lock cmpxchg` instruction. Support for locked instructions is immediate to add to our programming model.

Modulo these two changes, the instances $\mathcal{P}_A(I)$ coincide with the instrumentations $\mathcal{P}(I)_A$. Together with the argumentation in last two paragraphs this justifies the following theorem.

Theorem 4. *A parameterized program \mathcal{P} is not robust iff there is an attack A so that an instance $\mathcal{P}_A(I)$ of program \mathcal{P}_A reaches a goal configuration under SC.*

Reachability of a goal configuration in one instance of \mathcal{P}_A can be reformulated as a coverability problem for Petri nets, which is known to be decidable [26]. The key observation in the reduction to Petri nets is that threads in instance programs never use their identifiers, simply because they are copies of the same source code. This means there is no need to track the identity of threads, it is sufficient to count how many instances of a thread are in each state — a technique known as counter abstraction [13].

Theorem 5. *Robustness for parameterized programs over finite data domains is decidable and EXPSpace-hard — already for Boolean programs.*

For the lower bound, we in turn encode the coverability problem for Petri nets into robustness for parameterized programs [1,23]

7 Fence Insertion

To ease the presentation, we return to parallel programs. Since the algorithm only relies on a robustness checker, it carries over to the parametric setting.

Our goal is to insert a set of fences that ensure robustness of the resulting program. By *inserting a fence at label l* we mean the following modification of the program. Introduce a fresh label l_f . Then, translate each instruction `l: inst; goto l' ; into l_f : inst; goto l' ;`. Finally, add an instruction `l: mfence; goto l_f ;`.

We call a set of labels \mathcal{F} in program \mathcal{P} a *valid fence set* if inserting fences at these labels yields a robust program. We say that \mathcal{F} is *irreducible* if no strict subset is a valid fence set. In general, however, we would like to compute a valid fence set which is *optimal* in some sense. We pose the *fence computation problem*:

Given: A program \mathcal{P} and a strictly positive *cost function* $\mathcal{C}: \text{LAB} \rightarrow \mathbb{R}^+$.

Problem: Compute a valid fence set \mathcal{F} with $\sum_{l \in \mathcal{F}} \mathcal{C}(l)$ minimal.

Since we assume \mathcal{C} to be strictly positive, every optimal fence set is irreducible.

We consider two criteria of optimality: minimization of program size and maximization of program performance. By solving the problem for $\mathcal{C} \equiv 1$ we

compute a fence set of minimal size, thus minimizing the code size of the fenced program. Maximization of program performance requires minimizing the number of times memory fence instructions are executed: practical measurements [1] show that it is impossible to save CPU cycles by executing more fences, but with less stores in the TSO buffer. For this, $\mathcal{C}(l)$ is defined as the frequency at which instructions labeled by l occur in executions of the original program \mathcal{P} . Concrete values of \mathcal{C} can be either estimated by profiling or computed by mathematical reasoning about the program.

From a complexity point of view, fence computation is at least as hard as robustness. Indeed, robustness holds if and only if the optimal valid fence set is $\mathcal{F} = \emptyset$. Actually, since fence sets can be enumerated, computing an optimal valid fence set does not require more space than checking robustness. Notice that this also holds in the parameterized case.

Theorem 6. *For programs over finite domains, fence computation is PSPACE-complete. In the parameterized case, it is decidable and EXPSpace-hard.*

In the remainder of the section, we give a practical algorithm for computing optimal valid fence sets.

7.1 Fence Sets for Attacks

We say that a label l is *involved in the attack* $A = (t_A, \text{stinst}, \text{ldinst})$ if it belongs to some path in the control flow graph of t_A from the destination label of stinst to the source label of ldinst . We denote the set of all such labels by \mathcal{L}_A .

We call a set of labels \mathcal{F}_A an *eliminating fence set for attack A* if adding fences at all labels in \mathcal{F}_A eliminates the attack. Dekker’s algorithm has two eliminating fence sets: $\mathcal{F}_A = \{l_1\}$ eliminates the only attack by t_1 , and $\mathcal{F}_{A'} = \{l'_1\}$ eliminates the only attack by t_2 . Actually, the sets are *irreducible*: no strict subset eliminates the attack. Note that any irreducible eliminating set \mathcal{F}_A satisfies $\mathcal{F}_A \subseteq \mathcal{L}_A$.

Lemma 4. *Every irreducible valid fence set \mathcal{F} can be represented as a union of irreducible eliminating fence sets for all feasible attacks.*

Proof. By Theorem 1, fence set \mathcal{F} eliminates all feasible attacks. Therefore, it includes some irreducible eliminating fence set \mathcal{F}_A for every feasible attack A . By irreducibility, \mathcal{F} cannot contain labels outside the union of these \mathcal{F}_A sets. \square

In compliance with the lemma, in the Dekker’s program $\mathcal{F} = \mathcal{F}_A \cup \mathcal{F}_{A'}$.

Lemma 4 is useful for fence computation since optimal fence sets are always irreducible. All irreducible eliminating fence sets for attacks can be constructed by an exhaustive search through all selections of labels involved in the attack. For each candidate fence set, to judge whether it eliminates the attack, we check SC reachability in the instrumented program as described in Section 5.

Note that this search may raise an exponential number of reachability queries. In practice this rarely constitutes a problem. First, attacks seldom have a large number of involved labels, so the number of candidates is small. Second, the reachability checks can be avoided if a candidate fence set covers all the ways in the control flow graph from stinst to ldinst .

7.2 Computing an Optimal Valid Fence Set

To choose among the sets \mathcal{F}_A , we set up a 0/1-integer linear programming (ILP) problem $M_{\mathcal{P}} \cdot x_{\mathcal{P}} \geq b_{\mathcal{P}}$. The optimal solutions $f(x_{\mathcal{P}}) \rightarrow \min$ correspond to optimal fence sets. Here, 0/1 means the variables are restricted to yield Booleans.

We define inequalities that encode the feasible attacks with their corrections. Consider attack A for which we have determined the irreducible eliminating fence sets $\mathcal{F}_1, \dots, \mathcal{F}_n$. For each set, we introduce a variable $x_{\mathcal{F}_i}$ and set up Inequality (14)(left). It selects a fence set to eliminate the attack.

$$\sum_{1 \leq i \leq n} x_{\mathcal{F}_i} \geq 1 \quad \sum_{l \in \mathcal{F}_i} x_l \geq |\mathcal{F}_i| x_{\mathcal{F}_i} \quad f(x_{\mathcal{P}}) := \sum_{l \in \text{LAB}} \mathcal{C}(l) x_l. \quad (14)$$

When \mathcal{F}_i has been chosen, we insert a fence at each of its labels l . We add further variables x_l , and encode this insertion by Inequality (14)(center). By definition of the ILP, the variables $x_{\mathcal{F}_i}$ and x_l will only take Boolean values 0 or 1. So if $x_{\mathcal{F}_i}$ is set to 1, the inequality requires that all x_l with $l \in \mathcal{F}_i$ are set to 1.

Our goal is to select fences with minimal costs. We encode this into the objective function (14)(right). An optimal solution x^* of the resulting 0/1-ILP denotes the fence set $\mathcal{F}(x^*) := \{l \in \text{LAB} \mid x_l^* = 1\}$.

Theorem 7. $\mathcal{F}(x^*)$ is valid and optimal, and thus solves fence computation.

8 Experimental Evaluation

We implemented our algorithms in a prototype called TRENCHER [1]. The tool performs the reduction of robustness to SC reachability given in Section 5 and computes a minimal fence set as described in Section 7. TRENCHER executes independent reachability queries in parallel and uses Spin [16] as back-end model checker. With TRENCHER, we have performed a series of experiments.

8.1 Examples

The first class of examples are mutual exclusion protocols that are implemented via shared variables. These protocols are typically not robust under TSO and require additional fences after stores to synchronization variables. We studied robust and non-robust instances of Dekker and Peterson for two threads, as well as Lamport’s fast mutex [22] for three threads. Moreover, we checked the CLH and MCS Locks, robust list-based queue locks that use compare-and-set [15].

As second class of examples, we considered concurrent data structures. The Lock-Free Stack is a concurrent stack implementation using compare-and-swap [15]. Cilk’s THE WSQ is a work stealing queue from the implementation of the Cilk-5 programming language [12].

Finally, we consider miscellaneous concurrent algorithms that are known to be sensitive to program order relaxations. We analyse several instances of the Non-Blocking Write protocol [17]. NBWL is the spinlock + non-blocking write example considered by Owens in Section 8 of [24]. Finally, our tool discovers the known bug in Java’s Parker implementation that is due to TSO relaxations [11].

The test inputs are available online [1].

Program	T	L	I	RQ	NR1	NR2	R	F	Spin	Ver	Real
Peterson (non-robust)	2	14	18	23	2	12	9	2	7.7	0.5	2.9
Peterson (robust)	2	16	20	12	12	0	0	0	0.0	0.0	0.0
Dekker (non-robust)	2	24	30	95	12	28	55	4	31.7	2.1	14.2
Dekker (robust)	2	32	38	30	30	0	0	0	0.0	0.0	0.0
Lamport (non-robust)	3	33	36	36	9	15	12	6	14.4	6.0	5.9
Lamport (robust)	3	39	42	27	27	0	0	0	0.0	0.0	0.0
CLH Lock (robust)	7	62	58	54	48	6	0	0	4.9	0.2	1.6
MCS Lock (robust)	4	52	50	30	26	4	0	0	2.9	0.4	0.9
Lock-Free Stack (robust)	4	46	50	14	14	0	0	0	0.0	0.0	0.0
Cilk's THE WSQ (non-robust)	5	86	79	152	141	8	3	3	10.0	18.0	7.4
NBW2 (non-robust)	2	21	19	15	9	5	1	1	2.5	0.2	0.8
NBW3 (robust)	2	22	20	15	15	0	0	0	0.0	0.0	0.0
NBW4 (robust)	3	25	22	9	7	2	0	0	0.7	0.1	0.4
NBWL (robust)	4	45	45	30	26	4	0	0	2.7	0.2	0.7
Parker (non-robust)	2	9	8	2	0	1	1	1	0.5	0.0	0.3
Parker (robust)	2	10	9	2	2	0	0	0	0.0	0.0	0.0

Table 2: Benchmarking results.

8.2 Benchmarking

We executed TRENCHER on the examples, using a machine with Intel(R) Core(TM) i5 CPU M 560 @ 2.67GHz (4 cores) running GNU/Linux. Table 2 summarizes the results. The columns T, L, and I give the number of threads, labels, and instructions in the example. RQ is the number of reachability queries raised by TRENCHER. Provided the example is robust, this number is equal to the number of attacks (t_A , $stinst$, $ldinst$). NR1 is the number of verification queries that were answered negatively by TRENCHER itself, without running Spin. Such queries correspond to attacks where $stinst$ cannot be delayed past $ldinst$ because of memory fences or locked instructions in between. NR2 and R are the numbers of queries that are answered negatively/positively by the external model checker. Hence, $RQ = NR1 + NR2 + R$. F is the number of fences inserted.

The column Spin gives the total CPU time taken by Spin and Clang, the C compiler, to produce a verifier executable (pan). The column Ver provides the total CPU time taken by TRENCHER and the external verifier. Real is the wall-clock time in seconds of processing an example. All times are given in seconds.

8.3 Discussion

The analysis of robust algorithms is particularly fast. They typically only have a small number of attacks that have to be checked by a model checker. Robust Dekker and Peterson do not have such attacks at all. In the CLH and MCS locks, their number is less than 20%.

In some examples (non-robust Dekker, CLH Lock, NBW2, NBW4), up to 94% of the CPU time was spent on generating verifiers. This leaves room for improvement by switching to a model checker without compilation phase. For

some examples (LamNR, CLHLock), the wall-clock time constitutes 1/3 to 1/4 of the CPU time (4-cores). This confirms good parallelizability of the approach.

Remarkably, our trace-based analysis can establish robustness of the NBWL example, as opposed to the earlier analyses via triangular data races which would have to place a fence [24].

Acknowledgements

The second author has been granted by the Competence Center High Performance Computing and Visualization (CC-HPC) of the Fraunhofer Institute for Industrial Mathematics (ITWM).

References

1. Supplementary material. <http://concurrency.cs.uni-kl.de/trencher.html>.
2. P. A. Abdulla, M. F. Atig, Y.-F. Chen, C. Leonardsson, and A. Rezine. Counter-example guided fence insertion under tso. In *TACAS*, volume 7214 of *LNCS*, pages 204–219. Springer, 2012.
3. S. V. Adve and M. D. Hill. A unified formalization of four shared-memory models. *IEEE Trans. Parallel Distrib. Syst.*, 4(6):613–624, 1993.
4. J. Alglave. *A Shared Memory Poetics*. PhD thesis, University Paris 7, 2010.
5. J. Alglave and L. Maranget. Stability in weak memory models. In *CAV*, volume 6806 of *LNCS*, pages 50–66. Springer, 2011.
6. R. Alur, K. Mcmillan, and D. Peled. Model-checking of correctness conditions for concurrent objects. In *LICS*, pages 219–228. IEEE Computer Society Press, 1996.
7. M. F. Atig, A. Bouajjani, S. Burckhardt, and M. Musuvathi. On the verification problem for weak memory models. In *POPL*, pages 7–18. ACM, 2010.
8. A. Bouajjani, R. Meyer, and E. Möhlmann. Deciding robustness against total store ordering. In *ICALP*, volume 6756 of *LNCS*, pages 428–440. Springer, 2011.
9. S. Burckhardt and M. Musuvathi. Effective program verification for relaxed memory models. In *CAV*, volume 5123 of *LNCS*, pages 107–120. Springer, 2008.
10. J. Burnim, C. Stergiou, and K. Sen. Sound and complete monitoring of sequential consistency for relaxed memory models. In *TACAS*, volume 6605 of *LNCS*, pages 11–25. Springer, 2011.
11. D. Dice. A race in locksupport park() arising from weak memory models. https://blogs.oracle.com/dave/entry/a_race_in_locksupport_park, Nov 2009.
12. M. Frigo, C. E. Leiserson, and K.H. Randall. The implementation of the Cilk-5 multithreaded language. *SIGPLAN Not.*, 33(5):212–223, 1998.
13. S. M. German and P. A. Sistla. Reasoning about systems with many processes. *JACM*, 39:675–735, 1992.
14. P. B. Gibbons and E. Korach. Testing shared memories. *SIAM J. Comp.*, 26(4):1208–1244, 1997.
15. M. Herlihy and N. Shavit. *The Art of Multiprocessor Programming*. MKP, 2008.
16. G. J. Holzmann. The model checker SPIN. *IEEE Tr. Soft. Eng.*, 23:279–295, 1997.
17. H. Kopetz and J. Reisinger. The non-blocking write protocol NBW: A solution to a real-time synchronisation problem. In *IEEE Real-Time Systems Symposium*, pages 131–137. IEEE Computer Society Press, 1993.
18. D. Kozen. Lower bounds for natural proof systems. In *FOCS*, pages 254–266. IEEE Computer Society Press, 1977.
19. M. Kuperstein, M. T. Vechev, and E. Yahav. Partial-coherence abstractions for relaxed memory models. In *PLDI*, pages 187–198. ACM, 2011.

20. M. Kuperstein, M. T. Vechev, and E. Yahav. Automatic inference of memory fences. *SIGACT News*, 43(2):108–123, 2012.
21. L. Lamport. How to make a multiprocessor computer that correctly executes multiprocess programs. *IEEE Trans. Comp.*, 28(9):690–691, 1979.
22. L. Lamport. A fast mutual exclusion algorithm. *ACM Tr. Comp. Sys.*, 5(1), 1987.
23. R. Lipton. The reachability problem requires exponential space. Technical Report 62, Yale University, 1976.
24. S. Owens. Reasoning about the implementation of concurrency abstractions on x86-TSO. In *ECOOP*, volume 6183 of *LNCS*, pages 478–503. Springer, 2010.
25. S. Owens, S. Sarkar, and P. Sewell. A better x86 memory model: x86-TSO (extended version). Technical Report CL-TR-745, University of Cambridge, 2009.
26. C. Rackoff. The covering and boundedness problems for vector addition systems. *Theor. Comp. Sci.*, 6:223–231, 1978.
27. P. Sewell, S. Sarkar, S. Owens, F. Z. Nardelli, and M. O. Myreen. x86-TSO: a rigorous and usable programmer’s model for x86 multiprocessors. *CACM*, 53:89–97, 2010.
28. D. Shasha and M. Snir. Efficient and correct execution of parallel programs that share memory. *ACM TOPLAS*, 10(2):282–312, 1988.
29. V. Vafeiadis and F. Zappa Nardelli. Verifying fence elimination optimisations. In *SAS*, volume 6887 of *LNCS*, pages 146–162. Springer, 2011.

A Definition of Traces

Since the definition of traces in Section 3 was a bit brief, we recall here the full construction. Consider an SC or a TSO computation $\alpha \in \mathcal{C}_{\text{SC}}(\mathcal{P}) \cup \mathcal{C}_{\text{TSO}}(\mathcal{P})$. Its trace $\text{Tr}(\alpha)$ is a node-labelled graph $(N, \lambda, \rightarrow_{\text{po}}, \rightarrow_{\text{st}}, \rightarrow_{\text{src}})$ with nodes N , labelling $\lambda : N \rightarrow \text{ACT}$, and $\rightarrow_{\text{po}}, \rightarrow_{\text{st}}, \rightarrow_{\text{src}} \subseteq N \times N$ the aforementioned relations that define the edges. The *program order* is a union of $\rightarrow_{\text{po}} = \bigcup_{t \in \text{THRD}} \rightarrow_{\text{po}}^t$ of per thread total orders. The *store ordering* $\rightarrow_{\text{st}} = \bigcup_{a \in \text{DOM}} \rightarrow_{\text{st}}^a$ gives a total order for the stores to each address. We use the syntax $\text{max}(\rightarrow_{\text{po}}^t)$ and $\text{max}(\rightarrow_{\text{st}}^a)$ to access the maximal elements in these total orders. Finally, we have a *source relation* \rightarrow_{src} from stores to loads.

Traces are defined inductively, starting from the empty trace for the empty word ε . Assume we already constructed $\text{Tr}(\alpha) = (N, \lambda, \rightarrow_{\text{po}}, \rightarrow_{\text{st}}, \rightarrow_{\text{src}})$. In the definition of $\text{Tr}(\alpha \cdot \text{act}) := (N \cup \{n\}, \lambda', \rightarrow'_{\text{po}}, \rightarrow'_{\text{st}}, \rightarrow'_{\text{src}})$, the choice of n depends on the type of **act**. If we have a store, we use the moment the action was issued. Otherwise, we add a new node:

act = (t, st, a, v) Let n be the minimal node in $\rightarrow_{\text{po}}^t$ labelled by $\lambda(n) = \text{isu}$. We set $\lambda' := \lambda[n := \text{act}]$ and $\rightarrow'_{\text{po}} := \rightarrow_{\text{po}}$.
act $\neq (t, \text{st}, a, v)$ We add a fresh node $n \notin N$ to the trace, set $\lambda' := \lambda \cup \{(n, \text{act})\}$, and $\rightarrow'_{\text{po}} := \rightarrow_{\text{po}} \cup \{(\text{max}(\rightarrow_{\text{po}}^t), n)\}$.

The store order is updated only for stores (t, st, a, v) . We define $\rightarrow'_{\text{st}} := \rightarrow_{\text{st}} \cup \{(\text{max}(\rightarrow_{\text{st}}^a), n)\}$. The relation is not changed otherwise. The source relation is updated only for loads and stores. In case of a load (t, ld, a, v) we set $\rightarrow'_{\text{src}} := \rightarrow_{\text{src}} \cup \{(\text{max}(\rightarrow_{\text{st}}^a), n)\}$. In case of a store (t, st, a, v) we update the source relation for loads that read from the store early: for all nodes m with $n \rightarrow_{\text{po}}^+ m$ and $\lambda(m) = (t, \text{ld}, a, v)$ we set $\rightarrow'_{\text{src}} := (\rightarrow_{\text{src}} \setminus \{(*, m)\}) \cup \{(n, m)\}$.

Consider trace $\text{Tr}(\alpha)$ with $\alpha \in \mathcal{C}_{\text{TSO}}(\mathcal{P})$. The *conflict relation* \rightarrow_{cf} from load to store actions makes cyclic accesses in the trace visible. We define $\text{ld} \rightarrow_{\text{cf}} \text{st}$ if there is another store action st' in $\text{Tr}(\tau)$ that satisfies $\text{st}' \rightarrow_{\text{src}} \text{ld}$ and $\text{st}' \rightarrow_{\text{st}} \text{st}$. If ld reads the initial value of an address and st overwrites it, we also have $\text{ld} \rightarrow_{\text{cf}} \text{st}$. The *happens-before relation* of a trace is a union of all four relations: $\rightarrow_{\text{hb}} := \rightarrow_{\text{po}} \cup \rightarrow_{\text{st}} \cup \rightarrow_{\text{src}} \cup \rightarrow_{\text{cf}}$.

B Minimal Violations and Locality

in our earlier work [8] we showed that in a minimal violation only one thread reorders its actions. Since we employ here a more elaborate programming model, this locality result has to be checked again.

Consider a computation $\tau = \alpha \cdot a \cdot \beta \cdot b \cdot \gamma \in \mathcal{C}_{\text{TSO}}(\mathcal{P})$ with two actions a and b of the same thread $\text{thread}(a) = t = \text{thread}(b)$. We define the *distance* $d_\tau(a, b)$ between a and b in τ as the number of actions in β that also belong to this thread: $d_\tau(a, b) := \|\beta \downarrow t\|_{\text{len}}$. The *number of delays* $\#(\tau)$ in computation τ is the sum of distances between corresponding issue and store actions:

$$\#(\tau) := \sum_{\text{corr. isu, st in } \tau} d_\tau(\text{isu}, \text{st}).$$

We call a violating computation τ a *minimal violation* if it has a minimal number of delays among all violating computations. Clearly, a program \mathcal{P} has violating computations if and only if it has a minimal violation.

The following lemma says that if a store action has been delayed, then it has been delayed past a load action of the same thread. Moreover, the load did not read the value of this store action early.

Lemma 5. *Consider a minimal violation $\tau = \alpha \cdot \text{isu} \cdot \beta \cdot \text{st} \cdot \gamma \in C_{\text{TSO}}(\mathcal{P})$, where isu and st stem from the same instruction instance of thread t . Then $\beta \downarrow t$ is either empty, or $\beta \downarrow t = \beta' \cdot \text{ld} \cdot \beta''$ where ld is a load action with $\text{addr}(\text{ld}) \neq \text{addr}(\text{st})$ and β'' contains only store actions.*

Proof. Suppose β contains one or more actions of thread t . If all actions of thread t in β are stores, then also $\tau' = \alpha \cdot \beta \cdot \text{isu} \cdot \text{st} \cdot \gamma$ is a TSO computation of \mathcal{P} . It has the same trace as τ but $\#(\tau') < \#(\tau)$, which contradicts minimality of τ .

Otherwise let a be the last non-store action in $\beta \downarrow t$, i.e., $\beta = \beta_1 \cdot a \cdot \beta_2$ and all actions in β_2 are stores or belong to threads different from t . Since store actions cannot be delayed past a memory fence of the same thread, a is an issue action, a local action, or a load. In the former two cases, as well as if a is a load from $\text{addr}(\text{ld}) = \text{addr}(\text{st})$, delaying st past a can be avoided in the computation $\tau' = \alpha \cdot \text{isu} \cdot \beta_1 \cdot \beta_2 \cdot \text{st} \cdot a \cdot \gamma$ of \mathcal{P} . It has the same trace as τ and $\#(\tau') < \#(\tau)$, which contradicts minimality of τ . \square

In the remainder of the section, we develop a method to detect happens-before relations in a trace with the help of embedded computations. We relate two actions in a computation iff the corresponding nodes in the trace are related. To avoid case distinctions for issue and store actions that yield the same node in the trace, we introduce the *issue relation* \rightarrow_{isu} that links them: $\text{isu} \rightarrow_{\text{isu}} \text{st}$. We include \rightarrow_{isu} into \rightarrow_{hb} .

Definition 2 ([8]). *Let $\tau = \alpha \cdot a \cdot \beta \cdot b \cdot \gamma \in C_{\text{TSO}}(\mathcal{P})$. We say a happens-before b through β if there is a (potentially empty) subsequence $c_1 \dots c_n$ of β that satisfies (assuming $c_0 := a$ and $c_{n+1} := b$):*

$$a_i \rightarrow_{\text{hb}} a_{i+1} \quad \text{or} \quad a_i \rightarrow_{\text{po}}^+ a_{i+1} \quad \text{for all } i \in [0, n].$$

The next lemma states that the just defined relation is stable under insertion.

Lemma 6 ([8]). *Consider computations $\tau = \alpha \cdot a \cdot \beta \cdot b \cdot \gamma$ and $\tau' = \alpha' \cdot a \cdot \beta' \cdot b \cdot \gamma'$ in $C_{\text{TSO}}(\mathcal{P})$ so that $\tau \downarrow t = \tau' \downarrow t$ for every thread t . Let β be a subsequence of β' . Then if $a \rightarrow_{\text{hb}}^+ b$ through β then $a \rightarrow_{\text{hb}}^+ b$ through β' .*

The following lemma says that if two actions in a minimal violation are not related via $\rightarrow_{\text{hb}}^+$, they can be reordered without changing the trace and the order of actions within each thread.

Lemma 7 ([8]). *Consider a minimal violation $\tau = \alpha \cdot a \cdot \beta \cdot b \cdot \gamma \in C_{\text{TSO}}(\mathcal{P})$. Then (1) $a \rightarrow_{\text{hb}}^+ b$ through β or (2) there is $\tau' = \alpha \cdot \beta_1 \cdot b \cdot a \cdot \beta_2 \cdot \gamma \in C_{\text{TSO}}(\mathcal{P})$ so that $\text{Tr}(\tau) = \text{Tr}(\tau')$ and $\tau \downarrow t = \tau' \downarrow t$ for every thread t .*

Proof. We establish $\neg(1) \Rightarrow (2)$. Note that this proves the disjunction since $\neg(2) \Rightarrow (1)$ is the contrapositive. We proceed by induction on $\|\beta\|_{len}$ and slightly strengthen the hypothesis: we also show that β_2 is a subsequence of β .

Base case: $\|\beta\|_{len} = 0$. Then $\tau = \alpha \cdot a \cdot b \cdot \gamma$ and $a \not\rightarrow_{hb} b$. If $thread(a) = thread(b)$, then $b \rightarrow_{po}^+ a$. Therefore, b is a store action which has been delayed past a . Swapping a and b will save the delay without changing the trace, in contradiction to the minimality of τ .

If $thread(a) \neq thread(b)$, then either at least one of the two actions is local, the actions access different addresses, or both are loads. In all cases swapping them produces τ' as required in the statement of the lemma.

Step case: Assume the statement holds for $\|\beta\|_{len} \leq n$. Consider $\tau' = \alpha \cdot a \cdot \beta \cdot b \cdot \gamma$ with $\|\beta\|_{len} = n + 1$. Let c be the last action in $\beta = \beta' \cdot c$. Since $a \not\rightarrow_{hb}^+ b$ through β , then $a \not\rightarrow_{hb}^+ c$ through β' or $c \not\rightarrow_{hb} b$.

Let $a \not\rightarrow_{hb}^+ c$. We apply the induction hypothesis to τ with respect to a and c . This gives $\tau' = \alpha \cdot \beta'_1 \cdot c \cdot a \cdot \beta'_2 \cdot b \cdot \gamma$ with the same trace and thread computations as τ . Then, taking into account Lemma 6, we apply the hypothesis to τ' with respect to a and b . This yields $\tau'' = \alpha \cdot \beta'_1 \cdot c \cdot \beta'_{21} \cdot b \cdot a \cdot \beta'_{22} \cdot \gamma$ having the same trace and thread computations as τ' and τ . Note that β'_{22} is a subsequence of β'_2 , which in turn is a subsequence of β' and hence of β .

Let $c \not\rightarrow_{hb} b$. We apply the induction hypothesis to τ with respect to b and c , getting $\tau' = \alpha \cdot a \cdot \beta' \cdot b \cdot c \cdot \gamma$ with the same trace and thread computations as τ . Applying it again to τ' with respect to a and b gives $\tau'' = \alpha \cdot \beta'_1 \cdot b \cdot a \cdot \beta'_2 \cdot c \cdot \gamma$. The computation has the same trace and thread computations as τ' and τ . Since β'_2 is a subsequence of β' , $\beta'_2 \cdot c$ is a subsequence of β . \square

Lemma 8 (Locality [8]). *In a minimal violation, only a single thread delays stores.*

Proof. Consider a minimal violation $\tau \in C_{TSO}(\mathcal{P})$ and suppose at least two threads delayed stores. By Lemma 5, each store was delayed past a load of the same thread. Let st_2 of thread t_2 be the overall last delayed store in τ , and let ld_2 be the last load of t_2 overstepped by st_2 . Similarly, let st_1 be the overall last delayed store in a thread $t_1 \neq t_2$. Let ld_1 be the last load overstepped by st_1 .

The following fundamental mutual dispositions of reorderings are possible:

1. $\tau = \gamma_1 \cdot isu_1 \cdot \gamma_2 \cdot ld_1 \cdot \gamma_3 \cdot st_1 \cdot \gamma_4 \cdot isu_2 \cdot \gamma_5 \cdot ld_2 \cdot \gamma_6 \cdot st_2 \cdot \gamma_7$
2. $\tau = \gamma_1 \cdot isu_1 \cdot \gamma_2 \cdot ld_1 \cdot \gamma_3 \cdot isu_2 \cdot \gamma_4 \cdot ld_2 \cdot \gamma_5 \cdot st_2 \cdot \gamma_6 \cdot st_1 \cdot \gamma_7$
3. $\tau = \gamma_1 \cdot isu_1 \cdot \gamma_2 \cdot ld_1 \cdot \gamma_3 \cdot isu_2 \cdot \gamma_4 \cdot ld_2 \cdot \gamma_5 \cdot st_1 \cdot \gamma_6 \cdot st_2 \cdot \gamma_7$

In these three computations every pair (ld_i, st_i) provides a happens-before cycle: $st_i \rightarrow_{po}^+ ld_i$ and, by Lemma 7 and minimality, $ld_i \rightarrow_{hb}^+ st_i$ through the appropriate subrange of τ .

In the first disposition τ is not minimal, since it can be shortened to the violating computation $\tau' = \gamma_1 \cdot isu_1 \cdot \gamma_2 \cdot ld_1 \cdot \gamma_3 \cdot st_1 \cdot \beta$ with $\#(\tau') < \#(\tau)$. Here, β contains only store actions of t_2 that complete earlier issue actions.

In the second disposition τ is not minimal either. Starting from ld_1 , thread t_1 does not perform any actions, except delayed stores, until st_1 (Lemma 5).

Therefore, ld_1 and all program order later actions of t_1 can be safely removed from τ without affecting the happens-before cycle produced by t_2 . The resulting computation has a smaller number of delays (due to the removed ld_1), but its trace still includes the cycle by t_2 . A contradiction to minimality of τ .

Lastly, in the third case τ is also not minimal. First we delete γ_7 . Then we erase all actions from γ_6 that do not belong to t_2 : $\gamma'_6 = \gamma_6 \downarrow t_2$. By construction, the resulting computation τ' is a feasible TSO computation:

$$\tau' = \gamma_1 \cdot \text{isu}_1 \cdot \gamma_2 \cdot \text{ld}_1 \cdot \gamma_3 \cdot \text{isu}_2 \cdot \gamma_4 \cdot \text{ld}_2 \cdot \gamma_5 \cdot \text{st}_1 \cdot \gamma'_6 \cdot \text{st}_2.$$

Computation τ' still contains the happens-before cycle $\text{st}_1 \rightarrow_{\text{po}}^+ \text{ld}_1 \rightarrow_{\text{hb}}^+ \text{st}_1$ inherited from τ . Since deleting actions cannot increase the number of delays, $\#(\tau') = \#(\tau)$. Moreover, since τ is a minimal violation, so is τ' .

By Lemma 7, $\text{ld}_2 \rightarrow_{\text{hb}}^+ \text{st}_2$ through $\gamma_5 \cdot \text{st}_1 \cdot \gamma'_6$. By the choice of ld_1 and st_1 and in accordance with Lemma 5, $(\gamma_3 \cdot \text{isu}_2 \cdot \gamma_4 \cdot \text{ld}_2 \cdot \gamma_5) \downarrow t_1$ only contains delayed stores that were issued before ld_1 . By definition, γ'_6 does not contain actions of t_1 at all. Therefore, ld_1 is the program order last action of t_1 . It can be safely removed from τ' without affecting the cycle of t_2 . The resulting computation is

$$\tau'' = \gamma_1 \cdot \text{isu}_1 \cdot \gamma_2 \cdot \gamma_3 \cdot \text{isu}_2 \cdot \gamma_4 \cdot \text{ld}_2 \cdot \gamma_5 \cdot \text{st}_1 \cdot \gamma'_6 \cdot \text{st}_2.$$

Note that $\#(\tau'') < \#(\tau') = \#(\tau)$, but computation τ'' still contains the cycle $\text{st}_2 \rightarrow_{\text{po}}^+ \text{ld}_2 \rightarrow_{\text{hb}}^+ \text{st}_2$. A contradiction to minimality of τ . \square

C Soundness and Completeness of the Instrumentation

Theorem 8 (Soundness and Completeness). *Attack $A = (t_A, \text{stinst}, \text{ldinst})$ is feasible in program \mathcal{P} iff program \mathcal{P}_A reaches a goal configuration under SC.*

Proof. Soundness. Suppose the instrumented program reaches a goal configuration. For simplicity, assume that it immediately stops after this. Then the computation of the instrumented program looks like this:

$$\tau_A = \tau_1 \cdot \text{isu}_{\text{st}_A} \cdot \text{st}_A^{\text{aux}} \cdot \tau_2 \cdot \text{ld}_A \cdot \tau_3 \cdot \text{isu}_{\text{suc}} \cdot \text{st}_{\text{suc}}.$$

The last action, st_{suc} , is performed by a helper and sets variable `suc` to `true`, as required by the definition of goal configurations. This action originates from an instruction generated in accordance with (13). To reach this instruction, the helper has to enter its code copy.

As required by (8) and (9), for the first helper to enter its code copy, the attacker must set a `hb`-variable to a non-zero value by executing `ldinst` (action ld_A) instrumented in accordance with (2). For this, the attacker must enter its code copy and start performing stores to auxiliary addresses. Accordingly, the first attacker's store to an auxiliary address is denoted by st_A^{aux} in τ . It stems from the instrumented `stinst` (1) and is located before ld_A .

We elaborate on the contents of τ_1 , τ_2 , and τ_3 . First, the attacker and helpers execute the code of the original program (helpers — with an additional check at every instruction, (7)). In τ_2 the helpers continue to execute the code of the

original program. Shortly before performing ld_A and stopping, the attacker sets variable hb to true thus forcing the helpers to enter their code copies. Therefore all actions in τ_3 belong to helpers that have entered their code copies. Also, τ_2 only contains stores of the attacker to auxiliary addresses, and τ_3 does not contain attacker action at all, as follows from (2).

We now turn τ_A into the following TSO witness computation:

$$\tau = \tau'_1 \cdot \text{isu}_{\text{st}_A} \cdot \tau'_2 \cdot \text{ld}_A \cdot \tau'_3 \cdot \text{st}_A \cdot \tau'_4.$$

Here, τ'_1 is the subsequence of all τ_1 actions that are produced by instructions from \mathcal{P} (this is τ_1 without the conditionals introduced in (7)). Computation τ'_2 is the subsequence of all actions of τ_2 produced by instructions from \mathcal{P} and by their clones in the code copy of the attacker, except the store actions to auxiliary address. These store actions constitute τ'_4 . Finally, τ'_3 is the subsequence of all helper actions of τ_3 produced by clones of instructions from \mathcal{P} . We also strip the suffix d from the addresses of load and store actions in τ'_2 and τ'_4 .

That τ is a computation of program \mathcal{P} follows from the fact that τ_A is executable. We just removed actions produced by the instrumentation and replaced buffering by delaying of store actions; we did not change any data dependencies. The delaying of $\text{st}_A \cdot \tau'_4$ past ld_A is possible because the attacker did not execute memory fences between st_A^{aux} and ld_A , as guaranteed by (6).

Let us check that τ is a TSO witness (Figure 4). **(W1)** holds as in τ indeed only the attacker delays stores. The first delayed store st_A is an instance of stinst , load ld_A is an instance of ldinst and is the last action of the attacker that is overstepped by delayed stores, **(W2)** holds. For each act in $\text{ld}_A \cdot \tau_3 \cdot \text{st}_A$ it holds $\text{ld}_A \rightarrow_{\text{hb}}^* \text{act}$, **(W3)**. This is by construction of helpers in accordance with Lemma 3. Computation τ'_4 consists only of the stores delayed by the attacker, **(W4)**. **(W5)** holds due to the check in (2). So, τ is a TSO witness for attack A.

Completeness. Suppose there is a TSO witness τ for attack A as in Figure 4:

$$\tau = \tau_1 \cdot \text{isu}_{\text{st}_A} \cdot \tau_2 \cdot \text{ld}_A \cdot \tau_3 \cdot \text{st}_A \cdot \tau_4.$$

We show that the instrumented program has an execution that leads to a goal state. In the beginning, the instrumented attacker and helper threads execute instructions of the original program, namely those in τ_1 . The helpers actually execute these actions instrumented by (7), i.e., with an additional assert . These conditionals are executable because the attacker did not yet set variable hb .

Then the attacker executes $\llbracket \text{stinst} \rrbracket_{A1}$ (stinst is the instruction that produced isu_{st_A} in τ) and enters the code copy. Now all its stores will be executed on auxiliary addresses, as defined in (1) and (3). This means, they stay invisible to the helper threads as they were in the computation of the original program. Also, the instrumentation of loads (4) makes sure that they read buffered values, if they exist. Altogether this preserves the data dependencies from the original computation.

So the attacker executes the instructions that lie in τ_2 , instrumented by $\llbracket - \rrbracket_{A2}$. Note that τ_2 does not contain memory fences, otherwise st_A could not have

been delayed past ld_A in τ . Therefore, (6) cannot provoke a block of the attacker. The helpers still execute the actions of the original program, instrumented by (7). Finally, the attacker executes $ldinst$ which produced ld_A in τ , Equation (2). This is possible due to **(W5)**.

All actions in τ_3 belong to helpers. By **(W3)**, they are in happens-before relation with ld_A . Therefore, due to the instrumentation based on Lemma 3, the helpers are able to enter their code copies, (8) and (9), and execute the instructions that produced τ_3 . Note that the instrumentation of the code copy for helpers does not introduce any conditionals that could block the execution.

At least one of the helper’s actions in τ_3 performs a load or a store to the address used in st_A . Otherwise, **(W3)** would not hold (ld_A and the delayed write of the attacker use different addresses by **(W5)**). When performing the action in the instrumented program, the helper will set the hb -variable for the address used in st_A to a non-zero value, Equations (11) and (12). Therefore, at the next step the helper will be able to set suc to $true$ in accordance with (13) and make the instrumented program reach a goal state. \square

D Decidability and Complexity

The reductions of robustness to reachability and parameterized reachability are independent of the number of addresses and the structure of the data domain. Hence, without further assumptions the resulting reachability queries cannot be guaranteed to be decidable. We now discuss conditions on address space and data domain that render robustness decidable. Note that we only have to restrict these two dimensions. The instrumentation copes with the unbounded size store buffers. Moreover, we choose the verification technology so that it handles the unbounded number of threads required in parameterized reachability.

Parallel Programs with Finite Domains Consider a parallel program over a finite data domain, and hence finite address space. In this setting robustness is PSPACE-complete [8]. Our earlier proof is of complexity-theoretic nature: based on enumeration and not meant to be implemented. The instrumentation in this paper yields an alternative proof of membership in PSPACE that is conceptually simpler and allows us to reuse all techniques that have been developed for finite state verification.

Theorem 9. *Robustness for parallel programs over finite domains is PSPACE-complete.*

Parameterized Programs with Finite Domains Consider parameterized programs over finite domains. In this setting, decidability of robustness was open (our techniques from [8] do not carry over). With Theorem 4, we can now solve the problem and establish decidability. The key observation is that threads in instance programs never use their identifiers, simply because they are copies of the same source code. This means there is no need to track the identity of threads, it is sufficient to count how many instances of a thread are in each state — a technique known as counter abstraction [13]. Using this technique, we can reformulate the reachability problem for parameterized programs as a coverability problem for Petri nets. We briefly recall the basics on Petri nets.

Definitions A *Petri net* is a triple $N = (S, T, W)$ where S is a finite set of *places*, T is a finite set of *transitions* with $S \cap T = \emptyset$, and $W: (S \times T) \cup (T \times S) \rightarrow \mathbb{N}$ is a *weight function*. A *marking* is a function that assigns a natural number to each place: $M: S \rightarrow \mathbb{N}$. A *marked Petri net* is a pair (N, M_0) of a Petri net and an *initial marking* M_0 . A transition $t \in T$ is *enabled in marking* M if $M(\mathbf{s}) \geq W(\mathbf{s}, t)$ for all $\mathbf{s} \in S$. The *firing relation* $\langle \rangle \subseteq \mathbb{N}^{|S|} \times T \times \mathbb{N}^{|S|}$ contains a tuple (M_1, t, M_2) if transition t is enabled in M_1 and for all $\mathbf{s} \in S$ we have $M_2(\mathbf{s}) = M_1(\mathbf{s}) - W(\mathbf{s}, t) + W(t, \mathbf{s})$. We also write $M_1[t]M_2$. We extend the firing relation to sequences of transitions.

We say that a marking M is *reachable* in a marked Petri net (N, M_0) if there is a transition sequence $\sigma \in T^*$, such that $M_0[\sigma]M$. A marking M is *coverable* if there is a reachable marking M' so that $M(\mathbf{s})' \geq M(\mathbf{s})$ for all $\mathbf{s} \in S$.

Lemma 9 ([26]). *The problem to determine whether a marking M is coverable in a marked Petri net (N, M_0) is decidable.*

Reduction of parameterized reachability to Petri net coverability Let \mathcal{P} be a parameterized program with finite data domain DOM . We define a Petri net $N = (S, T, W)$ simulating the program.

For each pair of address and value $(a, v) \in \text{DOM} \times \text{DOM}$ we create a place $\mathbf{s}_{a,v}$. These places represent the state of the global memory: $M(\mathbf{s}_{a,v}) = 1$ corresponds to $\text{val}(a) = v$.

For each thread t_i that declares registers \bar{r}_i and has labels \bar{l}_i we create places $\mathbf{s}_{l,\bar{v}}$ for all $l \in \bar{l}_i$ and all $\bar{v} \in \text{DOM}^{|\bar{r}_i|}$. These places encode the number of thread instances in the given control state that have the given register valuation.

For each thread t_i we create a transition \mathbf{t}_i . Let $l_{0,i}$ be the initial label of t_i , \bar{r}_i be the registers declared by t_i , and $\bar{v}_{0,i}$ be a zero vector of length $|\bar{r}_i|$. Then we set $W(\mathbf{t}_i, \mathbf{s}_{l_{0,i}, \bar{v}_{0,i}}) = 1$. Transition \mathbf{t}_i effectively spawns an arbitrary number of copies of thread t_i that are all in the initial state.

Next we create transitions that simulate the instructions in each thread. We explain the construction for load instructions. The other instructions are handled along similar lines. Consider thread t_i with registers \bar{r}_i and a labelled load instruction $\text{linst} = l_1: r \leftarrow \text{mem}[f_a(\bar{r}_a)]; \text{goto } l_2; .$ For each value $v \in \text{DOM}$ and for each vector $\bar{v}_{reg} \in \text{DOM}^{|\bar{r}_i|}$ we create a transition $\mathbf{t} = \mathbf{t}_{\text{linst}, v, \bar{v}_{reg}}$. We set $W(\mathbf{s}_{l_1, \bar{v}_{reg}}, \mathbf{t}) = W(\mathbf{t}, \mathbf{s}_{l_2, \bar{v}'_{reg}}) = 1$ where $\bar{v}'_{reg} = \bar{v}_{reg}[r := v]$. Let $a = f_a(\bar{v}_{reg} \downarrow \bar{r}_a)$. Then we set $W(\mathbf{s}_{a,v}, \mathbf{t}) = W(\mathbf{t}, \mathbf{s}_{a,v}) = 1$. Transition \mathbf{t} is enabled if there is an instance of the thread in control state is l_1 so that its register valuation is \bar{v}_{reg} and address a being read holds value v . Firing the transition only updates the state of the thread instance: its program counter is set to label l_2 , and the value of register r is set to v .

We define the initial marking by $M_0(\mathbf{s}_{a,0}) = 1$ for all $a \in \text{DOM}$, and $M_0(\mathbf{s}) = 0$ for all other places $\mathbf{s} \in S$. Reaching a goal configuration $\text{val}(\text{suc}) = \text{true}$ in the parameterized program now corresponds to covering the following marking M_{suc} in the resulting Petri net: $M_{\text{suc}}(\mathbf{s}_{\text{suc}, \text{true}}) = 1$ and $M_{\text{suc}}(\mathbf{s}) = 0$ for all other places $\mathbf{s} \in S$. Combining this reduction with Lemma 9 gives Theorem 10.

Theorem 10. *Robustness for parameterized programs over finite domains is decidable.*

Lower bound The upper bound on robustness for parameterized programs depends on the data domain. Interestingly, an EXPSpace lower bound already holds for domains with two values. The proof reduces the coverability problem in Petri nets to robustness of parameterized programs. EXPSpace-hardness of coverability is a classic result by Lipton [23]. That we can restrict ourselves to domains with two elements means the control flow in a parameterized program is expressive enough to encode the Petri net behaviour.

The idea behind the construction is to take thread instances as tokens. Each thread has a label for each place in the Petri net, plus an additional label that indicates the token is currently not in use. The Petri net transitions are mimicked by a controller thread. It serialises the reading and writing of tokens, checks the coverability query, and then enters a non-robust situation. To read and write tokens, the controller communicates with the token threads via the memory. The construction requires locked instructions, which are immediate to add to our programming model.

Theorem 11. *Robustness for parameterized programs is EXPSpace-hard, for any domain with at least two elements.*