# Approximate Partial Order Reduction*

Chuchu Fan, Zhenqi Huang, and Sayan Mitra

University of Illinois at Urbana-Champaign, ECE Department
{cfan10, zhuang25, mitras}@illinois.edu

**Abstract.** We present a new partial order reduction method for reachability analysis of nondeterministic labeled transition systems over metric spaces. Nondeterminism arises from both the choice of the initial state and the choice of actions, and the number of executions to be explored grows exponentially with their length. We introduce a notion of $\varepsilon$-independence relation over actions that relates approximately commutative actions; $\varepsilon$-equivalent action sequences are obtained by swapping $\varepsilon$-independent consecutive action pairs. Our reachability algorithm generalizes individual executions to cover sets of executions that start from different, but $\delta$-close initial states, and follow different, but $\varepsilon$-independent, action sequences. The constructed over-approximations can be made arbitrarily precise by reducing the $\delta, \varepsilon$ parameters. Exploiting both the continuity of actions and their approximate independence, the algorithm can yield an exponential reduction in the number of executions explored. We illustrate this with experiments on consensus, platooning, and distributed control examples.

## 1 Introduction

Actions of different computing nodes interleave arbitrarily in distributed systems. The number of action sequences that have to be examined for state-space exploration grows exponentially with the number of nodes. Partial order reduction methods tackle this combinatorial explosion by eliminating executions that are *equivalent*, i.e., do not provide new information about reachable states (see [20,28,23] and the references therein). This equivalence is based on *independence* of actions: a pair of actions are independent if they commute, i.e., applying them in any order results in the same state. Thus, of all execution branches that start and end at the same state, but perform commuting actions in different order, only one has to be explored. Partial order reduction methods have become standard tools for practical software verification. They have been successfully applied to election protocols [2], indexers [19], file systems [9], security protocol [8], distributed schedulers [3], among many others.

Current partial order methods are limited when it comes to computation with numerical data and physical quantities (e.g., sensor networks, vehicle platoons, IoT applications, and distributed control and monitoring systems). First, a pair of actions are considered independent only if they commute exactly; actions that nearly commute—as are common in these applications—cannot be exploited for pruning the exploration.

* This work is supported by the grants CAREER 1054247 and CCF 1422798 from the National Science Foundation

Second, conventional partial order methods do not eliminate executions that start from nearly similar states and experience equivalent action sequences.

We address these limitations and propose a state space exploration method for nondeterministic, infinite state transition systems based on *approximate partial order reduction*. Our setup has two mild assumptions: (i) the state space of the transition system has a discrete part $L$ and a continuous part $X$ and the latter is equipped with a metric; (ii) the actions on $X$ are continuous functions. Nondeterminism arises from both the choice of the initial state and the choice of actions. Fixing an initial state $q_0$ and a sequence of actions $\tau$ (also called a *trace*), uniquely defines an execution of the system which we denote by $\xi_{q_0,\tau}$. For a given approximation parameter $\varepsilon \geq 0$, we define two actions $a$ and $b$ to be $\varepsilon$-independent if from any state $q$, the continuous parts of states resulting from applying action sequences $ab$ and $ba$ are $\varepsilon$-close. Two *traces* of $\mathcal{A}$ are $\varepsilon$-equivalent if they result from permuting $\varepsilon$-independent actions. To compute the reachable states of $\mathcal{A}$ using a finite (small) number of executions, the key is to generalize or expand an execution $\xi_{q_0,\tau}$ by a factor $r \geq 0$, so that, this expanded set contains all executions that start $\delta$-close to $q_0$ and experience action sequences that are $\varepsilon$-equivalent to $\tau$. We call this $r$ a $(\delta, \varepsilon)$-*trace equivalent discrepancy factor (*$ted$*)* for $\xi$.

For a fixed trace $\tau$, the only source of nondeterminism is the choice of the initial state. The reachable states from $B_\delta(q_0)$—a $\delta$-ball around $q_0$—can be over-approximated by expanding $\xi_{q_0,\tau}$ by a $(\delta, 0)$-$ted$. This is essentially the sensitivity of $\xi_{q_0,\tau}$ to $q_0$. Techniques for computing it are now well-developed for a broad class of models [13,11,14,16].

Fixing $q_0$, the only source of nondeterminism is the possible sequence of actions in $\tau$. The reachable states from $q_0$ following all possible valid traces can be over-approximated by expanding $\xi_{q_0,\tau}$ by a $(0, \varepsilon)$-$ted$, which includes states reachable by all $\varepsilon$-equivalent action sequences. Computing $(0, \varepsilon)$-$ted$ uses the principles of partial order reduction. However, unlike exact equivalence, here, starting from the same state, the states reached at the end of executing two $\varepsilon$-equivalent traces are not necessarily identical. This breaks a key assumption necessary for conventional partial order algorithms: here, an action enabled after $ab$ may not be enabled after $ba$. Of course, considering disabled actions can still give over-approximation of reachable states, but, we show that the precision of approximation can be improved arbitrarily by shrinking $\delta$ and $\varepsilon$.

Thus, the reachability analysis in this paper brings together two different ideas for handling nondeterminism: it combines sensitivity analysis with respect to initial state and $\varepsilon$-independence of actions in computing $(\delta, \varepsilon)$-$ted$, i.e., upper-bounds on the distance between executions starting from initial states that are $\delta$-close to each other and follow $\varepsilon$-equivalent action sequences (Theorem 1). As a matter of theoretical interest, we show that the approximation error can be made arbitrarily small by choosing sufficiently small $\delta$ and $\varepsilon$ (Theorem 2). We validate the correctness and effectiveness of the algorithm with three case studies where conventional partial order reduction would not help: an iterative consensus protocol, a simple vehicle platoon control system, and a distributed building heating system. In most cases, our reachability algorithm reduces the number of explored executions by a factor of $O(n!)$, for a time horizon of $n$, compared with exhaustive enumeration. Using these over-approximations, we could quickly decide safety verification questions. These examples illustrate that our method has the

potential to improve verification of a broader range of distributed systems for consensus [5,17,27,26], synchronization [31,29] and control [18,25].

*Related work.* There are two main classes of partial order reduction methods. The *persistent/ample set* methods compute a subset of enabled transitions –the persistent set (or ample set)– such that the omitted transitions are independent to those selected [10,2]. The reduced system which only considers the transitions in the persistent set is guaranteed to represent all behaviors of the original system. The persistent sets and the reduced systems are often derived by static analysis of the code. More recently, researchers have developed dynamic partial order reduction methods using the *sleep set* to avoid the static analysis [32,1,19]. These methods examine the history of actions taken by an execution and decide a set of actions that need to be explored in the future. The set of omitted actions is the sleep set. In [6], Cassez and Ziegler introduce a method to apply symbolic partial order reduction to infinite state discrete systems.

Analysis of sensitivity and the related notion of robustness analysis functions, automata, and executions has recently received significant attention [7,11,30]. Majumdar and Saha [24] present an algorithm to compute the output deviation with bounded disturbance combining symbolic execution and optimization. In [7] and [30], Chaudhuri etc., present algorithms for robustness analysis of programs and networked systems. Automatic techniques for local sensitivity analysis combining simulations and static analysis and their applications to verification of hybrid systems have been presented in [11,14,16].

In this paper, instead of conducting conventional partial order reduction, we propose a novel method of approximate partial order reduction, and combine it with sensitivity analysis for reachability analysis and safety verification for a broader class of systems.

## 2 Preliminaries

*Notations.* The state of our labeled transition system is defined by the valuations of a set of variables. Each variable $v$ has a type, $type(v)$, which is either the set of reals or some finite set. For a set of variables $V$, a valuation $\mathbf{v}$ maps each $v \in V$ to a point in $type(v)$. The set of all valuations of $V$ is $Val(V)$. $\mathbb{R}$ denotes the set of reals, $\mathbb{R}_{\geq 0}$ the set of non-negative reals, and $\mathbb{N}$ the set of natural numbers. For $n \in \mathbb{N}$, $[n] = \{0, \ldots, n-1\}$. The spectral radius $\rho(A)$ of a square matrix $A \in \mathbb{R}^{n \times n}$ is the largest absolute value of its eigenvalues. A square matrix $A$ is *stable* if its spectral radius $\rho(A) < 1$. For a set of tuples $S = \{\langle s_{j1}, \ldots, s_{jn} \rangle_j\}$, $S \lceil i$ denotes the set $\{s_{ji}\}$ which is the set obtained by taking the $i^{th}$ component of each tuple in $S$.

### 2.1 Transition systems

**Definition 1.** *A* labeled transition system $\mathcal{A}$ *is a tuple* $\langle X \cup L, \Theta, A, \rightarrow \rangle$ *where (i)* $X$ *is a set of* real-valued variables *and* $L$ *is a set of* finite-valued variables. $Q = Val(X \cup L)$ *is the* set of states*, (ii)* $\Theta \subseteq Q$ *is a set of* initial states *such that the sets of real-valued variables are compact, (iii)* $A$ *is a finite set of* actions*, and (iv)* $\rightarrow \subseteq Q \times A \times Q$ *is a* transition relation.

A state $q \in Q$ is a valuation of the real-valued and finite-valued variables. We denote by $q.X$ and $q.L$, respectively, the real-valued and discrete (finite-valued) parts of the state $q$. We will view the continuous part $q.X$ as a vector in $\mathbb{R}^{|X|}$ by fixing an arbitrary ordering of $X$. The norm $|\cdot|$ on $q.X$ is an arbitrary norm unless stated otherwise. For $\delta \geq 0$, the *$\delta$-neighborhood* of $q$ is denoted by $\mathcal{B}_\delta(q) \triangleq \{q' \in Q : q'.L = q.L \wedge |q'.X - q.X| \leq \delta\}$. For any $(q, a, q') \in \rightarrow$, we write $q \xrightarrow{a} q'$. For any action $a \in A$, its guard is the set $guard(a) = \{q \in Q \mid \exists q' \in Q, q \xrightarrow{a} q'\}$. We assume that guards are closed sets. An action $a$ is *deterministic* if for any state $q \in Q$, if there exists $q_1, q_2 \in Q$ with $q \xrightarrow{a} q_1$ and $q \xrightarrow{a} q_2$, then $q_1 = q_2$.

**Assumption 1** *(i) Actions are deterministic. For notational convenience, the name of an action $a$ is identified with its transition function, i.e., for each $q \in guard(a)$, $q \xrightarrow{a} a(q)$. We extend this notation to all states, i.e., even those outside $guard(a)$. (ii) For any state pair $q, q'$, if $q.L = q'.L$ then $a(q).L = a(q').L$.*

*Executions and traces.* For a deterministic transition system, a state $q_0 \in Q$ and a finite action sequence (also called a *trace*) $\tau = a_0 a_1 \dots a_{n-1}$ uniquely specifies a *potential execution* $\xi_{q_0,\tau} = q_0, a_0, q_1, a_1, \dots, a_{n-1}, q_n$ where for each $i \in [n]$, $a_i(q_i) = q_{i+1}$. A *valid execution* (also called execution for brevity) is a potential execution with (i) $q_0 \in \Theta$ and (ii) for each $i \in [n]$, $q_i \in guard(a_i)$. That is, a valid execution is a potential execution starting from the initial set with each action $a_i$ enabled at state $q_i$. For any potential execution $\xi_{q_0,\tau}$, its *trace* is the action sequence $\tau$, i.e., $trace(\xi_{q_0,\tau}) = \tau \in A^*$. We denote by $len(\tau)$ the length of $\tau$. For any for $i \in [len(\tau)]$, $\tau(i)$ is the $i$-th action in $\tau$. The length of $\xi_{q_0,\tau}$ is the length of its trace and $\xi_{q_0,\tau}(i) = q_i$ is the state visited after the $i$-th transition. The first and last state on a execution $\xi$ are denoted as $\xi.\mathsf{fstate} = \xi(0)$ and $\xi.\mathsf{lstate} = \xi(len(\xi))$.

For a subset of initial states $S \subseteq \Theta$ and a time bound $T \geq 0$, $\mathsf{Execs}(S, T)$ is the set of length $T$ executions starting from $S$. We denote the *reach set at time $T$* by $\mathsf{Reach}(S, T) \triangleq \{\xi.\mathsf{lstate} \mid \xi \in \mathsf{Execs}(S, T)\}$. Our goal is to precisely over-approximate $\mathsf{Reach}(\Theta, T)$ exploiting partial order reduction.

---

**automaton** Consensus($n \in \mathbb{N}, N \in \mathbb{N}$)
  **variables**
    $x : \mathbb{R}^n$
    $d : \mathbb{B}^N$
  **initially**
    $x[i] \in [-4, 4]$ **for each** $i \in [n]$
    $d[i] := false$ **for each** $i \in [n]$

**transitions**
  $\mathsf{a}_i$ **for each** $i \in [N]$
    **pre** $\neg d_i$
    **eff** $x := A_i x \wedge d[i] := true$
  $\mathsf{a}_\perp$
    **pre** $\wedge_{i \in [N]} d[i]$
    **eff** $d[i] := false$ **for each** $i \in [N]$

Fig. 1: Labeled transition system model of iterative consensus.

*Example 1 (Iterative consensus).* An $n$-dimensional iterative consensus protocol with $N$ processes is shown in Figure 1. The real-valued part of state is a vector $x$ in $\mathbb{R}^n$ and each process $i$ changes the state by the linear transformation $x \leftarrow A_i x$. The system evolves in rounds: in each round, each process $i$ updates the state exactly once but in arbitrary order. The boolean vector $d$ marks the processes that have acted in a round. The set of actions is $\{a_i\}_{i \in [N]} \cup \{a_\perp\}$. For each $i \in [N]$, the action $a_i$ is enabled when $d[i]$ is $false$ and when it occurs $x$ is updated as $A_i x$, where $A_i$ is an $n \times n$ matrix. The action

$a_\perp$ can occur only when all $d[i]$'s are set to $true$ and it resets all the $d[i]$'s to $false$. For an instance with $N = 3$, a valid execution could have the trace $\tau = a_0a_2a_1a_\perp a_1a_0a_2a_\perp$. It can be checked that Assumption 1 holds. In fact, the assumption will continue to hold if $A_ix$ is replaced by a nonlinear transition function $a_i : \mathbb{R}^n \to \mathbb{R}^n$.

### 2.2 Discrepancy functions

A discrepancy function bounds the changes in a system's executions as a continuous function of the changes in its inputs. Methods for computing discrepancy of dynamical and hybrid systems are now well-developed [22,14,12]. We extend the notion naturally to labeled transition systems: a discrepancy for an action bounds the changes in the continuous state brought about by its transition function.

**Definition 2.** *For an action $a \in A$, a continuous function $\beta_a : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a* discrepancy function *if for any pair of states $q, q' \in Q$ with $q.L = q'.L$, (i) $|a(q).X - a(q').X| \leq \beta_a(|q.X - q'.X|)$, and (ii) $\beta_a(\cdot) \to 0$ as $|q.X - q'.X| \to 0$.*

Property (i) gives an upper-bound on the changes brought about by action $a$ and (ii) ensures that the bound given by $\beta_a$ can be made arbitrarily precise. If the action $a$ is Lipschitz continuous with Lipschitz constant $L_a$, then $\beta_a(|q.X - q'.X|) = L_a(|q.X - q'.X|)$ can be used as a discrepancy function. Note that we do not assume the system is stable. As the following proposition states, given discrepancy functions for actions, we can reason about distance between executions that share the same trace but have different initial states.

**Proposition 1.** *Suppose each action $a \in A$ has a discrepancy function $\beta_a$. For any $T \geq 0$ and action sequence $\tau = a_0a_1a_2 \ldots a_T$, and for any pair of states $q, q' \in Q$ with $q.L = q'.L$, the last states of the pair of potential executions satisfy:*

$$\xi_{q,\tau}.\mathsf{lstate}.L = \xi_{q',\tau}.\mathsf{lstate}.L, \tag{1}$$

$$|\xi_{q,\tau}.\mathsf{lstate}.X - \xi_{q',\tau}.\mathsf{lstate}.X| \leq \beta_{a_T}\beta_{a_{T-1}} \ldots \beta_{a_0}(|q.X - q'.X|). \tag{2}$$

*Example 2.* Consider an instance of Consensus of Example 1 with $n = 3$ and $N = 3$ with the standard 2-norm on $\mathbb{R}^3$. Let the matrices $A_i$ be

$$A_0 = \begin{bmatrix} 0.2 & -0.2 & -0.3 \\ -0.2 & 0.2 & -0.1 \\ -0.3 & -0.1 & 0.3 \end{bmatrix}, A_1 = \begin{bmatrix} 0.2 & 0.3 & 0.2 \\ 0.3 & -0.2 & 0.3 \\ 0.2 & 0.3 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} -0.1 & 0 & 0.4 \\ 0 & 0.4 & -0.2 \\ 0.4 & -0.2 & -0.1 \end{bmatrix}.$$

It can be checked that for any pair $q, q' \in Q$ with $q.L = q'.L$, $|a_i(q).X - a_i(q').X|_2 \leq |A_i|_2|q.X - q'.X|_2$. Where the induced 2-norms of the matrices are $|A_0|_2 = 0.57, |A_1|_2 = 0.56, |A_2|_2 = 0.53$. Thus, for any $v \in \mathbb{R}_{\geq 0}$, we can use discrepancy functions for $a_0, a_1, a_2$: $\beta_{a_0}(v) = 0.57v, \beta_{a_1}(v) = 0.56v,$ and $\beta_{a_2}(v) = 0.53v$.

For actions with nonlinear transition functions, computing global discrepancy functions is difficult in general but local approaches using the eigenvalues of the Jacobian matrices are adequate for computing reachable sets from compact initial sets [16,21].

### 2.3 Combining sets of discrepancy functions

For a finite set of discrepancy functions $\{\beta_a\}_{a\in A'}$ corresponding to a set of actions $A' \subseteq A$, we define $\beta_{max} = \max_{a\in A'}\{\beta_a\}$ as $\beta_{max}(v) = \max_{a\in A'} \beta_a(v)$, for each $v \geq 0$. From Definition 2, for each $a \in S$, $\beta_a(|q.X - q'.X|) \to 0$ as $|q.X - q'.X| \to 0$. Hence, as the maximum of $\beta_a$, we have $\beta_{max}(|q.X - q'.X|) \to 0$ as $|q.X - q'.X| \to 0$. It can be checked that $\beta_{max}$ is a discrepancy function of each $a \in S$.

For $n \geq 0$ and a function $\beta_{max}$ defined as above, we define a function $\gamma_n = \sum_{i=0}^{n} \beta_{max}^i$; here $\beta^i = \beta \circ \beta^{i-1}$ for $i \geq 1$ and $\beta^0$ is the identity mapping. Using the properties of discrepancy functions as in Definition 2, we can show the following properties of $\{\gamma_n\}_{n\in\mathbb{N}}$.

**Proposition 2.** *Fix a finite set of discrepancy functions $\{\beta_a\}_{a\in A'}$ with $A' \subseteq A$. Let $\beta_{max} = \max_{a\in A'}\{\beta_a\}$. For any $n \geq 0$, $\gamma_n = \sum_{i=0}^{n} \beta_{max}^i$ satisfies (i) $\forall\, \varepsilon \in \mathbb{R}_{\geq 0}$ and any $n \geq n' \geq 0$, $\gamma_n(\varepsilon) \geq \gamma_{n'}(\varepsilon)$, and (ii) $\lim_{\varepsilon\to 0} \gamma_n(\varepsilon) = 0$.*

*Proof.* $(i)$ For any $n \geq 1$, we have $\gamma_n - \gamma_{n-1} = \beta_{max}^n$. Since $\beta_{max}^n = \max_{a\in S}\{\beta_a\}$ for some finite $S$, using Definition 2, $\beta_{max}^n$ takes only non-negative values. Hence, the sequence of functions $\{\gamma_n\}_{n\in\mathbb{R}_{\geq 0}}$ is non-decreasing.
$(ii)$ Using the property of discrepancy functions, we have $\lim_{\varepsilon\to 0} \beta_{max}(\varepsilon) = 0$. By induction on the nested functions, we have $\lim_{\varepsilon\to 0} \beta_{max}^i(0)$ for any $i \geq 0$. Hence for any $n \in \mathbb{R}_{\geq 0}$, $\lim_{\varepsilon\to 0} \gamma_n(\varepsilon) = \lim_{\varepsilon\to 0} \sum_{i=0}^{n} \beta_{max}^i(\varepsilon) = 0$.

The function $\gamma_n$ depends on the set of $\{\beta_a\}_{a\in A'}$, but as the $\beta$s will be fixed and clear from context, we write $\gamma_n$ for brevity.

## 3 Independent actions and neighboring executions

Central to partial order methods is the notion of independent actions. A pair of actions are independent if from any state, the occurrence of the two actions, in either order, results in the same state. We extend this notion and define a pair of actions to be $\varepsilon$-independent (Definition 3), for some $\varepsilon > 0$, if the continuous states resulting from swapped action sequences are within $\varepsilon$ distance.

### 3.1 Approximately independent actions

**Definition 3.** *For $\varepsilon \geq 0$, two distinct actions $a, b \in A$ are* $\varepsilon$-independent*, denoted by $a \overset{\varepsilon}{\sim} b$, if for any state $q \in Q$ (i) (Commutativity) $ab(q).L = ba(q).L$, and (ii) (Closeness) $|ab(q).X - ba(q).X| \leq \varepsilon$.*

The parameter $\varepsilon$ captures the degree of the approximation. Smaller the value of $\varepsilon$, more restrictive the independent relation. If $a$ and $b$ are $\varepsilon$-independent with $\varepsilon = 0$, then $ab(q) = ba(q)$ and the actions are independent in the standard sense (see e.g. Definition 8.3 of [4]). Definition 3 extends the standard definition in two ways. First, $b$ need not be enabled at state $a(q)$, and vice versa. That is, if $\xi_{q_0,ab}$ is an execution, we can only infer that $\xi_{q_0,ba}$ is a potential execution and not necessarily an execution. Secondly, with $\varepsilon > 0$, the continuous states can mismatch by $\varepsilon$ when $\varepsilon$-independent actions are

swapped. Consequently, an action $c$ may be enabled at $ab(q)$ but not at $ba(q)$. If $\xi_{q_0,abc}$ is a valid execution, we can only infer that $\xi_{q_0,bac}$ is a potential execution and not necessarily an execution.

We assume that the parameter $\varepsilon$ does not depend on the state $q$. When computing the value of $\varepsilon$ for concrete systems, we could first find an invariant for the state's real-valued variable $q.X$ such that $q.X$ is bounded, then find an upper-bound of $|ab(q).X - ba(q).X|$ as $\varepsilon$. For example, if $a$ and $b$ are both linear mappings with $a(q).X = A_1q.X + b_1$ and $b(q).X = A_2q.X + b_2$ and there is an invariant for $q.X$ is such that $|q.X| \leq r$, then it can be checked that $|ab(q).X - ba(q).X| = |(A_2A_1 - A_1A_2)q.X + (A_2b_1 - A_1b_2 + b_2 - b_1)| \leq |A_2A_1 - A_1A_2|r + |A_2b_1 - A_1b_2 + b_2 - b_1|$.

For a trace $\tau \in A^*$ and an action $a \in A$, $\tau$ is $\varepsilon$-independent to $a$, written as $\tau \overset{\varepsilon}{\sim} a$, if $\tau$ is empty string or for every $i \in [len(\tau)]$, $\tau(i) \overset{\varepsilon}{\sim} a$. It is clear that the approximate independence relation over $A$ is symmetric, but not necessarily transitive.

*Example 3.* Consider approximate independence of actions in Consensus. Fix any $i, j \in [N]$ such that $i \neq j$ and any state $q \in Q$. It can be checked that: $a_ia_j(q).d[k] = a_ja_i(q).d[k] = true$ *if* $k \in \{i, j\}$, otherwise it is $q.d[k]$. Hence, we have $a_ia_j(q).d = a_ja_i(q).d$ and the commutativity condition of Definition 3 holds. For the closeness condition, we have $|a_ia_j(q).x - a_ja_i(q).x|_2 = |(A_iA_j - A_jA_i)q.x|_2 \leq |A_iA_j - A_jA_i|_2|q.x|_2$. If the matrices $A_i$ and $A_j$ commute, then $a_i$ and $a_j$ are $\varepsilon$-approximately independent with $\varepsilon = 0$.

Suppose initially $x \in [-4, 4]^3$ then the 2-norm of the initial state is bounded by the value $4\sqrt{3}$. The specific matrices $A_i, i \in [3]$ presented in Example 2 are all stable, so $|a_i(q).x|_2 \leq |q.x|_2$, for each $i \in [3]$ and the norm of state is non-increasing in any transitions. Therefore, $Inv = \{x \in \mathbb{R}^3 : |x|_2 \leq 4\sqrt{3}\}$ is an invariant of the system. Together, we have $|a_0a_1(q).x - a_1a_0(q).x|_2 \leq 0.1$, $|a_0a_2(q).x - a_2a_0(q).x|_2 \leq 0.07$, and $|a_1a_2(q).x - a_2a_1(q).x|_2 \leq 0.17$. Thus, with $\varepsilon = 0.1$, it follows that $a_0 \overset{\varepsilon}{\sim} a_1$ and $a_0 \overset{\varepsilon}{\sim} a_2$ and $\overset{\varepsilon}{\sim}$ is not transitive, but with $\varepsilon = 0.2$, $\overset{\varepsilon}{\sim}$ is transitive.

### 3.2 $(\delta, \varepsilon)$-trace equivalent discrepancy for action pairs

Definition 3 implies that from a single state $q$, executing two $\varepsilon$-independent actions in either order, we end up in states that are within $\varepsilon$ distance. The following proposition uses discrepancy to bound the distance between states reached after performing $\varepsilon$-independent actions starting from *different* initial states $q$ and $q'$.

**Proposition 3.** *If a pair of actions $a, b \in A$ are $\varepsilon$-independent, and the two states $q, q' \in Q$ satisfy $q.L = q'.L$, then we have (i) $ba(q).L = ab(q').L$, and (ii) $|ba(q).X - ab(q').X| \leq \beta_b \circ \beta_a(|q.X - q'.X|) + \varepsilon$, where $\beta_a, \beta_b$ are discrepancy functions of $a, b$ respectively.*

*Proof.* Fix a pair of states $q, q' \in Q$ with $q.L = q'.L$. Since $a \overset{\varepsilon}{\sim} b$, we have $ba(q).L = ab(q).L$. Using the Assumption, we have $ab(q).L = ab(q').L$. Using triangular inequality, we have $|ba(q).X - ab(q').X| \leq |ba(q).X - ba(q').X| + |ba(q').X - ab(q').X|$. The first term is bounded by $\beta_b \circ \beta_a(|q.X - q'.X|)$ using Proposition 1 and the second is bounded by $\varepsilon$ by Definition 3, and hence, the result follows.

## 4 Effect of $\varepsilon$-independent traces

In this section, we will develop an analog of Proposition 3 for $\varepsilon$-independent traces (action sequences) acting on neighboring states.

### 4.1 $\varepsilon$-equivalent traces

First, we define what it means for two finite traces in $A^*$ to be $\varepsilon$-equivalent.

**Definition 4.** *For any $\varepsilon \geq 0$, we define a relation $R \subseteq A^* \times A^*$ such that $\tau R \tau'$ iff there exists $\sigma, \eta \in A^*$ and $a, b \in A$ such that $a \overset{\varepsilon}{\sim} b$, $\tau = \sigma ab\eta$, and $\tau' = \sigma ba\eta$. We define an equivalence relation $\overset{\varepsilon}{\equiv} \subseteq A^* \times A^*$ called* $\varepsilon$-equivalence*, as the reflexive and transitive closure of $R$.*

That is, two traces $\tau, \tau' \in A^*$ are $\varepsilon$-equivalent if we can construct $\tau'$ from $\tau$ by performing a sequence of swaps of consecutive $\varepsilon$-independent actions.

In the following proposition, states that the last states of two potential executions starting from the same initial discrete state (location) and resulting from equivalent traces have identical locations.

**Proposition 4.** *Fix potential executions $\xi = \xi_{q_0,\tau}$ and $\xi' = \xi_{q_0',\tau'}$. If $q_0.L = q_0'.L$ and $\tau \overset{\varepsilon}{\equiv} \tau'$, then $\xi.\mathsf{lstate}.L = \xi'.\mathsf{lstate}.L$.*

*Proof.* If $\tau = \tau'$, then the proposition follows from the Assumption. Suppose $\tau \neq \tau'$, from Definition 4, there exists a sequence of action sequences $\tau_0, \tau_1, \ldots, \tau_k$ to join $\tau$ and $\tau'$ by swapping neighboring approximately independent actions. Precisely the sequence $\{\tau_i\}_{i=0}^k$ satisfies: (i) $\tau_0 = \tau$ and $\tau_k = \tau'$, and (ii) for each pair $\tau_i$ and $\tau_{i+1}$, there exists $\sigma, \eta \in A^*$ and $a, b \in A$ such that $a \overset{\varepsilon}{\sim} b$, $\tau_i = \sigma ab\eta$, and $\tau_{i+1} = \sigma ba\eta$. From Definition 3, swapping approximately independent actions preserves the value of the discrete part of the final state. Hence for any $i \in [k]$, $\xi_{q_0,\tau_i}.\mathsf{lstate}.L = \xi_{q_0,\tau_{i+1}}.\mathsf{lstate}.L$. Therefore, $\xi.\mathsf{lstate}.L = \xi'.\mathsf{lstate}.L$.

Next, we relate pairs of potential executions that result from $\varepsilon$-equivalent traces and initial states that are $\delta$-close.

**Definition 5.** *Given $\delta, \varepsilon \geq 0$, a pair of initial states $q_0, q_0'$, and a pair traces $\tau, \tau' \in A^*$, the corresponding potential executions $\xi = \xi_{q_0,\tau}$ and $\xi' = \xi_{q_0',\tau'}$ are* $(\delta, \varepsilon)$-related*, denoted by $\xi \overset{\delta,\varepsilon}{\approx} \xi'$, if $q_0.L = q_0'.L$, $|q_0.X - q_0'.X| \leq \delta$, and $\tau \overset{\varepsilon}{\equiv} \tau'$.*

*Example 4.* In Example 3, we show that $a_0 \overset{\varepsilon}{\sim} a_1$ and $a_0 \overset{\varepsilon}{\sim} a_2$ with $\varepsilon = 0.1$. Consider the executions $\xi = q_0, a_0, q_1, a_1, q_2, a_2, q_3, a_\perp, q_4$ and $\xi' = q_0', a_1, q_1', a_2, q_2', a_0, q_3', a_\perp, q_4'$. with traces $trace(\xi) = a_0a_1a_2a_\perp$ and $trace(\xi') = a_1a_2a_0a_\perp$. For $\varepsilon = 0.1$, we have $a_0a_1a_2a_\perp \overset{\varepsilon}{\equiv} a_1a_0a_2a_\perp$ and $a_1a_0a_2a_\perp \overset{\varepsilon}{\equiv} a_1a_2a_0a_\perp$. Since the equivalence relation $\overset{\varepsilon}{\equiv}$ is transitive, we have $trace(\xi) \overset{\varepsilon}{\equiv} trace(\xi')$. Suppose $q_0 \in \mathcal{B}_\delta(q_0')$, then $\xi$ and $\xi'$ are $(\delta, \varepsilon)$-related executions with $\varepsilon = 0.1$.

It follows from Proposition 4 that the discrete state (locations) reached by any pair of $(\delta, \varepsilon)$-related potential executions are the same. At the end of this section, in Lemma 2, we will bound the distance between the continuous state reached by $(\delta, \varepsilon)$-related potential executions. We define in the following this bound as what we call *trace equivalent discrepancy factor* $(ted)$, which is a constant number that works for all possible values of the variables starting from the initial set. Looking ahead, by bloating a single potential execution by the corresponding $ted$, we can over-approximate the reachset of all related potential executions. This will be the basis for the reachability analysis in Section 5.

**Definition 6.** *For any potential execution $\xi$ and constants $\delta, \varepsilon \geq 0$, a* $(\delta, \varepsilon)$-trace equivalent discrepancy factor $(ted)$ *is a nonnegative constant $r \geq 0$, such that for any $(\delta, \varepsilon)$-related potential finite execution $\xi'$,*

$$|\xi'.\mathsf{lstate}.X - \xi.\mathsf{lstate}.X| \leq r.$$

That is, if $r$ is a $(\delta, \varepsilon)$-$ted$, then the $r$-neighborhood of $\xi$'s last state $\mathcal{B}_r(\xi.\mathsf{lstate})$ contains the last states of all other $(\delta, \varepsilon)$-related potential executions.

### 4.2 $(0, \varepsilon)$-trace equivalent discrepancy for traces (on the same initial states)

In this section, we will develop an inductive method for computing $(\delta, \varepsilon)$-$ted$. We begin by bounding the distance between potential executions that differ only in the position of a single action.

**Lemma 1.** *Consider any $\varepsilon \geq 0$, an initial state $q_0 \in Q$, an action $a \in A$ and a trace $\tau \in A^*$ with $len(\tau) \geq 1$. If $\tau \overset{\varepsilon}{\sim} a$, then the potential executions $\xi = \xi_{q_0,\tau a}$ and $\xi' = \xi_{q_0,a\tau}$ satisfy*

*(i)* $\xi'.\mathsf{lstate}.L = \xi.\mathsf{lstate}.L$ *and*
*(ii)* $|\xi'.\mathsf{lstate}.X - \xi.\mathsf{lstate}.X| \leq \gamma_{n-1}(\varepsilon)$*, where $\gamma_n$ corresponds to the set of discrepancy functions $\{\beta_c\}_{c\in\tau}$ for the actions in $\tau$.*

*Proof.* Part (i) directly follows from Proposition 4. We will prove part (ii) by induction on the length of $\tau$.
**Base**: For any trace $\tau$ of length 1, $\xi$ and $\xi'$ are of the form $\xi = q_0, b_0, q_1, a, q_2$ and $\xi' = q_0, a, q_1', b_0, q_2'$. Since $a \overset{\varepsilon}{\sim} b_0$ and the two executions start from the same state, it follows from Definition 3 that $|q_2'.X - q_2.X| \leq \varepsilon$. Recall from the preliminary that $\gamma_0(\varepsilon) = \beta^0(\varepsilon) = \varepsilon$. Hence $|q_2'.X - q_2.X| \leq \gamma_0(\varepsilon)$ holds for trace $\tau$ with $len(\tau) = 1$.
**Induction**: Suppose the lemma holds for any $\tau$ with length at most $n-1$. Fixed any $\tau = b_0 b_1 \dots b_{n-1}$ of length $n$, we will show the lemma holds for $\tau$.

Let the potential executions $\xi = \xi_{q_0,\tau a}$ and $\xi' = \xi_{q_0,a\tau}$ be the form

$$\xi = q_0, b_0, q_1, b_1, ..., b_{n-1}, q_n, a, q_{n+1},$$
$$\xi' = q_0, a, q_1', b_0, q_2', b_1, ..., b_{n-1}, q_{n+1}'.$$

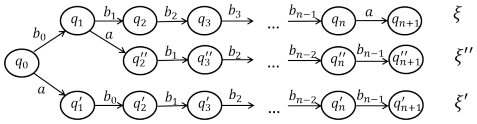

Fig. 2: Potential executions $\xi$, $\xi'$, and $\xi''$.

It suffices to prove that $|\xi.\mathsf{lstate}.X - \xi'.\mathsf{lstate}.X| = |q_{n+1}.X - q'_{n+1}.X| \leq \gamma_{n-1}(\varepsilon)$. We first construct a potential execution $\xi'' = \xi_{q_0,b_0ab_1\ldots b_{n-1}}$ by swapping the first two actions of $\xi'$. Then, $\xi''$ is of the form: $\xi'' = q_0, b_0, q_1, a, q''_2, b_1, ..., b_{n-1}, q''_{n+1}$. The potential executions $\xi, \xi'$ and $\xi''$ are shown in Figure 2. We first compare the potential executions $\xi$ and $\xi''$. Notice that, $\xi$ and $\xi''$ share a common prefix $q_0, b_0, q_1$. Starting from $q_1$, the action sequence of $\xi''$ is derived from $trace(\xi)$ by inserting action $a$ in front of the action sequence $\tau' = b_1 b_2 \ldots b_{n-1}$.

Since $\tau' \overset{\varepsilon}{\sim} a$, applying the induction hypothesis on the length $n-1$ action sequence $\tau'$, we get $|q_{n+1}.X - q''_{n+1}.X| \leq \gamma_{n-2}(\varepsilon)$. Then, we compare the potential executions $\xi'$ and $\xi''$. Since $b_0 \overset{\varepsilon}{\sim} a$, by applying the property of Definition 3 to the first two actions of $\xi'$ and $\xi''$, we have $|q'_2.X - q''_2.X| \leq \varepsilon$. We note that $\xi'$ and $\xi''$ have the same suffix of action sequence from $q'_2$ and $q''_2$. Using Proposition 1 from states $q'_2$ and $q''_2$, we have

$$|q'_{n+1}.X - q''_{n+1}.X| \leq \beta_{b_1}\beta_{b_2}\ldots\beta_{b_{n-1}}(|q'_2.X - q''_2.X|) \leq \beta^{n-1}(\varepsilon). \quad (3)$$

Combining the bound on $|q'_2.X - q''_2.X|$ and (3) with triangular inequality, we have $|q_{n+1}.X - q'_{n+1}.X| \leq |q_{n+1}.X - q''_{n+1}.X| + |q'_{n+1}.X - q''_{n+1}.X| \leq \gamma_{n-2}(\varepsilon) + \beta^{n-1}(\varepsilon) = \gamma_{n-1}(\varepsilon)$.

### 4.3 $(\delta, \varepsilon)$-trace equivalent discrepancy for traces

Lemma 1 gives a way to compute $(0, \varepsilon)$-$ted$. Now, we generalize this to compute $(\delta, \varepsilon)$-$ted$, for $(\delta, \varepsilon)$-related potential executions, with any $\delta \geq 0$. The following lemma gives an inductive way of constructing $ted$ as an action $a$ is appended to a trace $\tau$. The proof is analog to the proof of Lemma 1 and is provided in the full version of this paper [15].

**Lemma 2.** *For any potential execution $\xi = \xi_{q_0,\tau}$ and constants $\delta, \varepsilon \geq 0$, if $r$ is a $(\delta, \varepsilon)$-$ted$ for $\xi$, and the action $a \in A$ satisfies $\tau \overset{\varepsilon}{\sim} a$, then $r' = \beta_a(r) + \gamma_{len(\tau)-1}(\varepsilon)$ is a $(\delta, \varepsilon)$-$ted$ for $\xi_{q_0,\tau a}$.*

## 5 Reachability with approximate partial order reduction

We will present our main algorithm (Algorithm 2) for reachability analysis with approximate partial order reduction in this section. The core idea is to over-approximate $\mathsf{Reach}(B_\delta(q_0), T)$ by (a) computing the actual execution $\xi_{q_0,\tau}$ and (b) expanding this $\xi_{q_0,\tau}$ by a $(\delta, \varepsilon)$-$ted$ to cover all the states reachable from any other $(\delta_0, \varepsilon)$-related potential execution. Combining such over-approximations from a cover of $\Theta$, we get over-approximations of $\mathsf{Reach}(\Theta, T)$, and therefore, Algorithm 2 can be used to soundly check for bounded safety or invariance. The over-approximations can be made arbitrarily precise by shrinking $\delta_0$ and $\varepsilon$ (Theorem 2). Of course, at $\varepsilon = 0$ only traces that are exactly equivalent to $\tau$ will be covered, and nothing else. Algorithm 2 avoids computing $(\delta_0, \varepsilon)$-related executions, and therefore, gains (possibly exponential) speedup.

The key subroutine in Algorithm 2 is $CompTed$ which computes the $ted$ by adding one more action to the traces. It turns out that, the $ted$ is independent of $q_0$, but only depends on the sequence of actions in $\tau$. $CompTed$ is used to compute $\delta_t$ from $\delta_{t-1}$, such that, $\delta_t$ is the $ted$ for the length $t$ prefix of $\xi$. Let action $a$ be the $t^{th}$ action and $\xi = \xi_{q_0,\tau a}$. If $a$ is $\varepsilon$-independent to $\tau$, then the $ted$ $\delta_t$ can be computed from $\delta_{t-1}$ just

using Lemma 2. For the case where $a$ is not $\varepsilon$-independent to the whole sequence $\tau$, we would still want to compute a set of executions that $\xi_{q_0,\tau a}$ can cover. We observe that, with appropriate computation of $ted$, $\xi_{q_0,\tau a}$ can cover all executions of the form $\xi_{q_0,\phi a\eta}$, where $\phi a\eta$ is $\varepsilon$-equivalent to $\tau a$ and $a \notin \eta$. In what follows, we introduce this notion of *earliest equivalent position of $a$ in $\tau$* (Definition 7), which is the basis for the $CompTed$ subroutine, which in turn is then used in the main reachability Algorithm 2.

### 5.1 Earliest equivalent position of an action in a trace

For any trace $\tau \in A^*$ and action $a \in \tau$, we define $lastPos(\tau, a)$ as the largest index $k$ such that $\tau(k) = a$. The earliest equivalent position, $eep(\tau, a, \varepsilon)$ is the minimum of $lastPos(\tau', a)$ in any $\tau'$ that is $\varepsilon$-equivalent to $\tau a$.

**Definition 7.** *For any trace $\tau \in A^*$, $a \in A$, and $\varepsilon > 0$, the* earliest equivalent position *of $a$ on $\tau$ is $eep(\tau, a, \varepsilon) \triangleq \min_{\tau' \overset{\varepsilon}{\equiv} \tau a} lastPos(\tau', a)$.*

For any trace $\tau a$, its $\varepsilon$-equivalent traces can be derived by swapping consecutive $\varepsilon$-independent action pairs. Hence, the eep of $a$ is the leftmost position it can be swapped to, starting from the end. Any equivalent trace of $\tau a$ is of the form $\phi a\eta$ where $\phi$ and $\eta$ are the prefix and suffix of the last occurrence of action $a$. Hence, equivalently: $eep(\tau, a, \epsilon) = \min_{\phi a\eta \overset{\varepsilon}{\equiv} \tau a,\ a \notin \eta} len(\phi)$. In the full version of this paper [15] we give a simple $O(len(\tau)^2)$ algorithm for computing $eep()$. If the $\varepsilon$-independence relation is symmetric, then it $eep$ can be computed in $O(len(\tau))$ time.

*Example 5.* In Example 3, we showed that $a_0 \overset{\varepsilon}{\sim} a_1$ and $a_0 \overset{\varepsilon}{\sim} a_2$ with $\varepsilon = 0.1$; $a_\perp$ is not $\varepsilon$-independent to any actions. What is $eep(a_\perp a_0 a_1, a_2, \varepsilon)$? We can swap $a_2$ ahead following the sequence $\tau a_2 = a_\perp a_0 a_1 a_2 \overset{\varepsilon}{\equiv} a_\perp a_1 a_0 a_2 \overset{\varepsilon}{\equiv} a_\perp a_1 a_2 a_0$. As $a_\perp$ and $a_1$ are not independent of $a_2$, it cannot occur earlier. $eep(a_\perp a_0 a_1, a_2, \varepsilon) = 2$.

### 5.2 Reachability using $(\delta, \varepsilon)$-trace equivalent discrepancy

$CompTed$ (Algorithm 1) takes inputs of trace $\tau$, a new action to be added $a$, a parameter $r \geq 0$ such that $r$ is a $(\delta_0, \varepsilon)$-$ted$ for the potential execution $\xi_{q_0,\tau}$ for some initial state $q_0$, initial set radius $\delta_0$, approximation parameter $\varepsilon \geq 0$, and a set of discrepancy functions $\{\beta_a\}_{a \in A}$. It returns a $(\delta_0, \varepsilon)$-$ted$ $r'$ for the potential execution $\xi_{q_0,\tau a}$.

**Algorithm 1** $CompTed(\tau, a, r, \varepsilon, \{\beta_a\}_{a \in A})$

1: $\beta \leftarrow \max_{b \in \tau a}\{\beta_b\}$; $k \leftarrow eep(\tau, a, \epsilon)$; $t \leftarrow len(\tau)$;
2: **if** $k = t$ **then** $r' \leftarrow \beta_a(r)$ **else** $r' \leftarrow \beta_a(r) + \gamma_{t-k-1}(\varepsilon)$
3: **return** $r'$;

**Lemma 3.** *For some initial state $q_0$ and initial set size $\delta_0$, if $r$ is a $(\delta_0, \varepsilon)$-ted for $\xi_{q_0,\tau}$ then value returned by $CompTed()$ is a $(\delta_0, \varepsilon)$-ted for $\xi_{q_0,\tau a}$.*

*Proof.* Let us fix some initial state $q_0$ and initial set size $\delta_0$.

Let $\xi_t = \xi_{q_0,\tau}$ be the potential execution starting from $q_0$ by taking the trace $\tau$, and $\xi_{t+1} = \xi_{q_0,\tau a}$. Fix any $\xi'$ that is $(\delta_0, \varepsilon)$-related to $\xi_{t+1}$. From Proposition 4, $\xi'.\mathsf{lstate}.L = \xi_{t+1}.\mathsf{lstate}.L$. It suffice to prove that $|\xi'.\mathsf{lstate}.X - \xi_{t+1}.\mathsf{lstate}.X| \leq r'$.

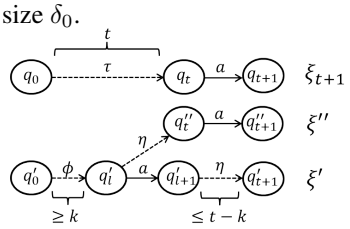

Fig. 3: Potential executions $\xi_{t+1}$, $\xi'$,$\xi''$

Since $trace(\xi') \overset{\varepsilon}{\equiv} \tau a$, action $a$ is in the sequence $trace(\xi')$. Partitioning $trace(\xi')$ on the last occurrence of $a$, we get $trace(\xi') = \phi a \eta$ for some $\phi, \eta \in A^*$ with $a \notin \eta$. Since $k$ is the $eep$, from Definition 7, the position of the last occurrence of $a$ on $trace(\xi')$ is at least $k$. Hence we have $len(\phi) \geq k$ and $len(\eta) = t - len(\phi) \leq t - k$. We construct another potential execution $\xi'' = \xi_{q'_0,\phi\eta a}$ with the same initial state as $\xi'$. The executions $\xi_{t+1}, \xi'$ and $\xi''$ are illustrated in Figure 3.

$q_t$ is the last state of the execution $\xi_t$. From the assumption, $\mathcal{B}_r(q_t)$ is an over-approximation of the reachset at step $t$. We note that the length $t$ prefix $\xi''$ is $(\delta_0, \varepsilon)$-related to $\xi_t$. Therefore, $|q_t.X - q''_t.X| \leq r$. Using the discrepancy function of action $a$, we have

$$|q_{t+1}.X - q''_{t+1}.X| \leq \beta_a(|q_t.X - q''_t.X|) \leq \beta_a(r). \tag{4}$$

We will quantify the distance between $\xi'$ and $\xi''$. There are two cases:
(i) If $k = t$ then, $len(\eta) \leq t - k = 0$, that is, $\eta$ is an empty string. Hence, $\xi'$ and $\xi''$ are indeed identical and $q'_{t+1} = q''_{t+1}$. Thus from (4), $|q_{t+1}.X - q'_{t+1}.X| = |q_{t+1}.X - q''_{t+1}.X| \leq \beta_a(r)$, and the lemma holds. (ii) Otherwise $k < t$ and from Lemma 1, we can bound the distance between $\xi'$ and $\xi''$ as $|q'_{t+1}.X - q''_{t+1}.X| \leq \gamma_{len(\eta)-1}(\varepsilon) \leq \gamma_{t-k-1}(\varepsilon)$. Combining with (4), we get $|q_{t+1}.X - q'_{t+1}.X| \leq |q_{t+1}.X - q''_{t+1}.X| + |q'_{t+1}.X - q''_{t+1}.X| \leq \beta_a(r) + \gamma_{t-k-1}(\varepsilon)$.

Next, we present the main reachability algorithm which uses $CompTed$. Algorithm 2 takes inputs of an initial set $\Theta$, time horizon $T$, two parameters $\delta_0, \varepsilon \geq 0$, and a set of discrepancy functions $\{\beta_a\}_{a \in A}$. It returns the over-approximation of the reach set for each time step.

The algorithm first computes a $\delta_0$-cover $Q_0$ of the initial set $\Theta$ such that $\Theta \subseteq \cup_{q_0 \in Q_0} \mathcal{B}_\delta(q_0)$ (Line 2). The **for**-loop from Line 3 to Line 14 will compute the over-approximation of the reachset from each initial cover $\mathsf{Reach}(\mathcal{B}_{\delta_0}(q_0), t)$. The over-approximation from each cover is represented as a collection $\langle R_0, \ldots, R_T \rangle$, where each $R_t$ is a set of tuples $\langle \tau_t, q_t, \delta_t \rangle$ such that (i) the traces $R_t \lceil 1$ and their $\varepsilon$-equivalent traces contain the traces of all valid executions of length $t$, (ii) the traces in $R_t \lceil 1$ are mutually non-$\varepsilon$-equivalent, (iii) for each tuple $\delta_t$ is the $(\delta_0, \varepsilon)$-$ted$ for $\xi_{q_0,\tau_t}$,

For each initial cover $\mathcal{B}_{\delta_0}(q_0)$, $R_0$ is initialized as the tuple of empty string, the initial state $q_0$ and size $\delta_0$ (Line 4). Then the reachset over-approximation is computed recursively for each time step by checking for the maximum set of enabled actions $EA$ for the set of states $\mathcal{B}_{\delta_t}(q_t)$ (Line 8), and try to attach each enabled action $a \in EA$ to $\tau_t$ unless $\tau_t a$ is $\varepsilon$-equivalent to some length $t+1$ trace that is already in $R_{t+1} \lceil 1$. This is where the major reduction happens using approximate partial order reduction.

If not, the $(\delta_0, \varepsilon)$-*ted* for $\xi_{q_0,\tau_t a}$ will be computed using $CompTed$, and new tuple $\langle \tau_t a, q_{t+1}, \delta_{t+1} \rangle$ will be added to $R_{t+1}$ (Line 13).

If there are $k$ actions in total and they are mutually $\varepsilon$-independent, then as long as the numbers of each action in $\tau_t$ and $\tau_t'$ are the same, $\tau_t \overset{\varepsilon}{\equiv} \tau_t'$. Therefore, in this case, $R_t$ contains at most $\binom{t+k-1}{k-1}$ tuples. Furthermore, for any length $t$ trace $\tau_t$, if all actions in $\tau_t$ are mutually $\varepsilon$-independent, the algorithm can reduce the number of executions explored by $O(t!)$. Essentially, each $\tau_t \in R_t \lceil 1$ is a representative trace for the length $t$ $\varepsilon$-equivalence class.

**Algorithm 2** Reachability algorithm to over-approximate $\mathsf{Reach}(\Theta, T)$

1: **Input**: $\Theta, T, \varepsilon, \delta_0, \{\beta_a\}$;
2: $Q_0 \leftarrow \delta_0$-$cover(\Theta)$; $\mathcal{R} \leftarrow \emptyset$
3: **for** $q_0 \in Q_0$ **do**
4: $\quad R_0 \leftarrow \{\langle '', q_0, \delta_0 \rangle\}$;
5: $\quad$ **for** $t = [T]$ **do**
6: $\quad\quad R_T \leftarrow \emptyset$;
7: $\quad\quad$ **for** each $\langle \tau_t, q_t, \delta_t \rangle \in R_t$ **do**
8: $\quad\quad\quad EA \leftarrow enabledactions(\mathcal{B}_{\delta_t}(q_t))$;
9: $\quad\quad\quad$ **for** $a \in EA$ **do**
10: $\quad\quad\quad\quad$ **if** $\forall \tau_{t+1} \in R_{t+1} \lceil 1, \neg \left(\tau_t a \overset{\varepsilon}{\equiv} \tau_{t+1}\right)$ **then**;
11: $\quad\quad\quad\quad\quad q_{t+1} \leftarrow a(q_t)$
12: $\quad\quad\quad\quad\quad \delta_{t+1} \leftarrow CompTed(\tau_t, a, \delta_t, \varepsilon, \{\beta_a\}_{a \in A})$
13: $\quad\quad\quad\quad\quad R_{t+1} \leftarrow R_{t+1} \cup \langle \tau_t a, q_{t+1}, \delta_{t+1} \rangle$
14: $\quad \mathcal{R} \leftarrow \mathcal{R} \cup \langle R_0, \ldots, R_T \rangle$
15: **return** $\mathcal{R}$;

Theorem 1 shows that Algorithm 1 indeed computes an over-approximation for the reachsets, and Theorem 2 states that the over-approximation can be made arbitrarily precise by reducing the size of $\delta_0, \varepsilon$.

**Theorem 1 (Soundness).** *Set $\mathcal{R}$ returned by Algorithm 2, satisfies $\forall t = 0, \ldots, T$,*

$$\mathsf{Reach}(\Theta, t) \subseteq \bigcup_{R_t \in \mathcal{R} \lceil t} \bigcup_{\langle \tau, q, \delta \rangle \in R_t} \mathcal{B}_\delta(q). \tag{5}$$

*Proof.* Since $\cup_{q_0 \in Q_0} \mathcal{B}_\delta(q_0) \supseteq \Theta$, it suffices to show that at each time step $t = 0, \ldots, T$, the $R_t$ computed in the **for**-loop from Line 4 to Line 13 satisfy $\mathsf{Reach}(\mathcal{B}_{\delta_0}(q_0), t) \subseteq \cup_{\langle \tau, q, \delta \rangle \in R_t} \mathcal{B}_\delta(q)$. Fix any $q_0 \in Q_0$, we will prove by induction.

Base case: initially before any action happens, the only valid trace is the empty string $''$ and the initial set is indeed $\mathcal{B}_{\delta_0}(q_0)$.

Induction step: assume that at time step $t < T$, the union of all the traces $R_t \lceil 1$ and their $\varepsilon$-equivalent traces contain the traces of all length $t$ valid executions, and for each tuple $\langle \tau_t, q_t, \delta_t \rangle \in R_t$, $\delta_t$ is a $(\delta_0, \varepsilon)$-*ted* for $\xi_{q_0,\tau_t}$. That is, $\mathcal{B}_{\delta_t}(q_t)$ contains the final states of all $(\delta_0, \varepsilon)$-related executions to $\xi_{q_0,\tau_t}$. This is sufficient for showing that $\mathsf{Reach}(\mathcal{B}_{\delta_0}(q_0), t) \subseteq \cup_{\langle \tau, q, \delta \rangle \in R_t} \mathcal{B}_\delta(q)$.

Since for each tuple contained in $R_t$, we will consider the maximum possible set of actions enabled at Line 8 and attempts to compute the $(\delta_0, \varepsilon)$-*ted* for $\xi_{q_0,\tau_t a}$. If $\tau_t a$ is not $\varepsilon$-equivalent to any of the length $t+1$ traces that has already been added to $R_{t+1}$, then Lemma 3 guarantees that the $q_{t+1}$ and $\delta_{t+1}$ computed at Line 11 and 12 satisfy that $\delta_{t+1}$ is the $(\delta_0, \varepsilon)$-*ted* for $\xi_{q_0,\tau_t a}$. Otherwise, $\tau_t a$ is $\varepsilon$-equivalent to some trace $\tau_{t+1}$ that has already been added to $R_{t+1}$, then for any initial state $q_0'$ that is $\delta_0$-close to $q_0$, $\xi_{q_0',\tau_t a}$ and $\xi_{q_0,\tau_{t+1}}$ are $(\delta_0, \varepsilon)$-related and the final state of $\xi_{q_0',\tau_t a}$ is already contained in $\mathcal{B}_{\delta_{t+1}}(q_{t+1})$. Therefore, the union of all the traces $R_{t+1} \restriction 1$ and their $\varepsilon$-equivalent traces contain the traces of all length $t+1$ valid executions, and for each tuple $\langle \tau_{t+1}, q_{t+1}, \delta_{t+1} \rangle \in R_{t+1}$, $\delta_{t+1}$ is a $(\delta_0, \varepsilon)$-*ted* for $\xi_{q_0,\tau_{t+1}}$, which means $\mathsf{Reach}(\mathcal{B}_{\delta_0}(q_0), t+1) \subseteq \cup_{\langle \tau, q, \delta \rangle \in R_{t+1}} \mathcal{B}_\delta(q)$. So the theorem holds.

**Theorem 2 (Precision).** *For any $r > 0$, there exist $\delta_0, \varepsilon > 0$ such that, the reachset over-approximation $\mathcal{R}$ computed by Algorithm 2 satisfies $\forall t = 0, \ldots, T$,*

$$\bigcup_{R_t \in \mathcal{R} \lceil t} \bigcup_{\langle \tau, q, \delta \rangle \in R_t} \mathcal{B}_\delta(q) \subseteq \mathcal{B}_r(\mathsf{Reach}(\Theta, t)). \tag{6}$$

The proof is based on the fact that $\delta_t$ computed in Algorithm 2 converges to zero as $\delta_0 \to 0$ and $\varepsilon \to 0$, and the details are given in the full version of the paper [15]. Notice that as $\delta_0$ and $\varepsilon$ go to 0, the Algorithm 2 actually converges to a simulation algorithm which simulates every valid execution from a single initial state.

## 6 Experimental evaluation of effectiveness

We discuss the results from evaluating Algorithm 2 in three case studies. Our Python implementation runs on a standard laptop (Intel Core™ i7-7600 U CPU, 16G RAM).

*Iterative consensus.* This is an instance of Consensus (Example 1) with 3 continuous variables and 3 actions $a_0, a_1, a_2$. We want to check if the continuous states converge to $[-0.4, 0.4]^3$ in 3 rounds starting from a radius 0.5 ball around $[2.5, 0.5, -3]$. Figure 4 (Left) shows reachset over-approximation computed and projected on $x[0]$. The blue and red curves give the bounds. As the figure shows, $x[0]$ converges to $[-0.4, 0.4]$ at round 3; and so do $x[1]$ and $x[2]$ (not shown). We also simulated 100 random valid executions (yellow curves) from the initial set and validate that indeed the over-approximation is sound.

Recall, three actions can occur in any order in each round, i.e., $3! = 6$ traces per round, and $6^3 = 216$ executions from a single initial state up to 3 rounds. We showed in Example 3 that $a_0 \overset{\varepsilon}{\sim} a_1$ and $a_0 \overset{\varepsilon}{\sim} a_2$ with $\varepsilon = 0.1$. Therefore, $a_0a_1a_2 \overset{\varepsilon}{\equiv} a_1a_0a_2 \overset{\varepsilon}{\equiv} a_1a_2a_0$ and $a_0a_2a_1 \overset{\varepsilon}{\equiv} a_2a_0a_1 \overset{\varepsilon}{\equiv} a_2a_1a_0$, and Algorithm 2 explored only 2 (length 12) executions from a set of initial states for computing the bounds. The running time for Algorithm 2 is 1 millisecond while exploring all valid executions from even only a single state took 20 milliseconds.

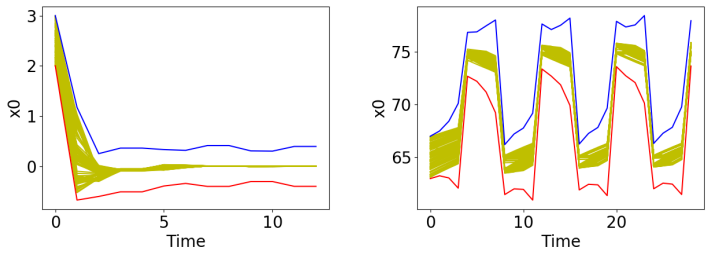


Fig. 4: Reachset computations. The blue curves are the upper bound of the reachsets and the red curves are the lower bound of the reachsets. Between the blue and red curves, the yellow curves are 100 random simulations of valid executions. *Left:* Linear transition system. *Right:* Room heating system.

*Platoon.* Consider an $N$ car platoon on a single lane (see Figure 7 in the the full version of the paper [15] for the pseudocode and details). Each car can choose one of three actions at each time step: $a$ (accelerate), $b$ (brake), or $c$ (cruise). Car 0 can choose any action at each time step; remaining cars try to keep safe distance with predecessor by choosing accelerate ($a$) if the distance is more than 50, brake ($b$) if the distance is less than 30, and cruise ($c$) otherwise.

Consider a 2-car platoon and a time horizon of $T = 10$. We want to verify that the cars maintain safe separation. Reachset over-approximations projected on the position variables are shown in Figure 5, with 100 random simulations of valid executions as a sanity check. Car 0 has lots of choices and it's position over-approximation diverges (Figure 5). Car 1's position depends on its initial relative distance with Car 0. It is also easy to conclude from Figure 5 that two cars maintain safe relative distance for these different initial states.

From a single initial state, in every step, Car 0 has 3 choices, and therefore there are $3^{10}$ possible executions. Considering a range of initial positions for two cars, there are infinitely many execution, and $9^{10}$ (around 206 trillion) possible traces. With $\epsilon = 0.282$, Algorithm 2 explored a maximum of $\binom{18}{8} = 43758$ traces; the concrete number varies for different initial sets. The running time for Algorithm 2 is 5.1 milliseconds while exploring all valid executions from even only a single state took 2.9 seconds.

For a 4-car platoon and a time horizon of $T = 10$, there are $81^{10}$ possible traces considering a range of initial positions. With $\varepsilon = 0.282$, Algorithm 2 explored 7986 traces to conclude that all cars maintain safe separation for the setting where all cars are initially separated by a distance of 40 and has an initial set radius of 4. The running time for Algorithm 2 is 62.3 milliseconds, while exploring all valid executions from even only a single state took 6.2 seconds.

*Building heating system.* Consider a building with $N$ rooms, each with a heater (see the full version of the paper [15] for pseudocode and details). For $i \in [N]$, $x[i] \in \mathbb{R}$ is the temperature of room $i$ and $m[i] \in \{0, 1\}$ captures the off/on state of it's heater. The controller measures the temperature of rooms periodically; based on these measurements ($y[i]$) heaters turn on or off. These decisions are made asynchronously across rooms in arbitrary order. The room temperature $x[i]$ changes linearly according to the heater input $m[i]$, the thermal capacity of the room, and the thermal coupling across adjacent rooms as given in the benchmark problem of [18]. For $i \in [N]$, actions $\mathsf{on}_i$, $\mathsf{off}_i$ capture the decision making process of room $i$ on whether or not to turn on the heater. Time

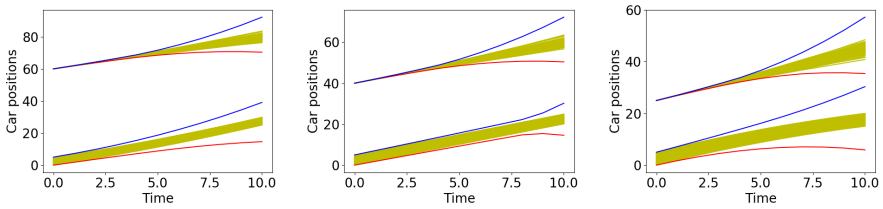


Fig. 5: Position over-approximations for 2 cars. The blue curves are the upper bound of the reachsets and the red curves are the lower bound of the reachsets. Between the blue and red curves, the yellow curves are 100 random simulations of valid executions. Car1's initial position is in the range $[0, 5]$, Car2's initial position is 60 (Left), 40 (Center) and 25 (Right).

elapse is captured by a flow action that updates the temperatures. We want to verify that the room temperatures remain in the $[60, 79]$ range.

Consider a building with $N = 3$ rooms. In the full version of the paper [15], we provide computation details to show that for any $i, j \in [3]$ with $i \neq j$, $a \in \{\mathsf{on}_i, \mathsf{off}_i\}$ and $b \in \{\mathsf{on}_j, \mathsf{off}_j\}$, $a \overset{\varepsilon}{\sim} b$ with $\varepsilon = 0.6$; but, flow is not independent with any other actions. Computed reachset over-approximation for 8 rounds and projected on the temperature of Room 0 is shown in Figure 4 (*Right*). Indeed, temperature of Room 0 is contained within the range.

For a round, where each room makes a decision once in arbitrary order, there are $3! = 6$ $\varepsilon$-equivalent action sequences. Therefore, from a single initial state, there are $6^8$ (1.6 million) valid executions. Algorithm 2, in this case explore only one (length 32) execution with $\varepsilon = 0.6$ to approximate all executions starting from an initial set with radius $\delta = 2$. The running time for Algorithm 2 is 1 millisecond while exploring all valid executions from even only a single state took 434 seconds.

## 7 Conclusion

We proposed a partial order reduction technique for reachability analysis of infinite state transition systems that exploits approximate independence and bounded sensitivity of actions to reduce the number of executions explored. This relies on a novel notion of $\varepsilon$-independence that generalizes the traditional notion of independence by allowing approximate commutation of actions. With this $\varepsilon$-independence relation, we have developed an algorithm for soundly over-approximating reachsets of all executions using only $\varepsilon$-equivalent traces. The over-approximation can also be made arbitrarily precise by reducing the size of $\delta, \varepsilon$. In experimental evaluation with three case studies we observe that it can reduce the number of executions explored exponentially compared to explicit computation of all executions.

The results suggest several future research directions. In Definition 3, $\varepsilon$-independent actions are required to be approximately commutative globally. For reachability analysis, this definition could be relaxed to actions that approximately commute locally over parts of the state space. An orthogonal direction is to apply this reduction technique to verify temporal logic properties and extend it to hybrid models.

## References


1. Abdulla, P., Aronis, S., Jonsson, B., Sagonas, K.: Optimal dynamic partial order reduction. In: ACM SIGPLAN Notices. vol. 49, pp. 373–384. ACM (2014)
2. Alur, R., Brayton, R.K., Henzinger, T.A., Qadeer, S., Rajamani, S.K.: Partial-order reduction in symbolic state space exploration. In: International Conference on Computer Aided Verification. pp. 340–351. Springer (1997)
3. Baier, C., Größer, M., Ciesinski, F.: Partial order reduction for probabilistic systems. In: QEST. vol. 4, pp. 230–239 (2004)
4. Baier, C., Katoen, J.P., Larsen, K.G.: Principles of model checking. MIT press (2008)
5. Blondel, V., Hendrickx, J.M., Olshevsky, A., Tsitsiklis, J., et al.: Convergence in multiagent coordination, consensus, and flocking. In: IEEE Conference on Decision and Control. vol. 44, p. 2996. IEEE; 1998 (2005)
6. Cassez, F., Ziegler, F.: Verification of concurrent programs using trace abstraction refinement. In: Logic for Programming, Artificial Intelligence, and Reasoning. pp. 233–248. Springer (2015)
7. Chaudhuri, S., Gulwani, S., Lublinerman, R.: Continuity and robustness of programs. Communications of the ACM 55(8), 107–115 (2012)
8. Clarke, E., Jha, S., Marrero, W.: Partial order reductions for security protocol verification. In: International Conference on Tools and Algorithms for the Construction and Analysis of Systems. pp. 503–518. Springer (2000)
9. Clarke, E.M., Grumberg, O., Minea, M., Peled, D.: State space reduction using partial order techniques. International Journal on Software Tools for Technology Transfer 2(3), 279–287 (1999)
10. Clarke, E.M., Grumberg, O., Peled, D.: Model checking. MIT press (1999)
11. Donzé, A.: Breach, a toolbox for verification and parameter synthesis of hybrid systems. In: Computer Aided Verification (CAV) (2010)
12. Donzé, A., Maler, O.: Systematic simulation using sensitivity analysis. In: Hybrid Systems: Computation and Control, pp. 174–189. Springer (2007)
13. Duggirala, P.S., Mitra, S., Viswanathan, M.: Verification of annotated models from executions. In: EMSOFT (2013)
14. Duggirala, P.S., Mitra, S., Viswanathan, M., Potok, M.: C2E2: A verification tool for stateflow models. In: Tools and Algorithms for the Construction and Analysis of Systems. Lecture Notes in Computer Science, vol. 9035, pp. 68–82. Springer Berlin Heidelberg (2015)
15. Fan, C., Huang, Z., Mitra, S.: Approximate partial order reduction (full version) (May 2018), https://arxiv.org/abs/1610.06317
16. Fan, C., Mitra, S.: Bounded verification with on-the-fly discrepancy computation. In: International Symposium on Automated Technology for Verification and Analysis. pp. 446–463. Springer (2015)
17. Fang, L., Antsaklis, P.J.: Information consensus of asynchronous discrete-time multi-agent systems. In: Proceedings of the 2005, American Control Conference, 2005. pp. 1883–1888. IEEE (2005)
18. Fehnker, A., Ivančić, F.: Benchmarks for hybrid systems verification. In: International Workshop on Hybrid Systems: Computation and Control. pp. 326–341. Springer (2004)
19. Flanagan, C., Godefroid, P.: Dynamic partial-order reduction for model checking software. In: ACM Sigplan Notices. vol. 40, pp. 110–121. ACM (2005)
20. Godefroid, P., van Leeuwen, J., Hartmanis, J., Goos, G., Wolper, P.: Partial-order methods for the verification of concurrent systems: an approach to the state-explosion problem, vol. 1032. Springer Heidelberg (1996)

21. Huang, Z., Fan, C., Mereacre, A., Mitra, S., Kwiatkowska, M.: Simulation-based verification of cardiac pacemakers with guaranteed coverage. IEEE Design & Test 32(5), 27–34 (Oct 2015)
22. Huang, Z., Mitra, S.: Proofs from simulations and modular annotations. In: Proceedings of the 17th international conference on Hybrid systems: computation and control. pp. 183–192. ACM (2014)
23. Kurshan, R., Levin, V., Minea, M., Peled, D., Yenigün, H.: Static partial order reduction. In: International Conference on Tools and Algorithms for the Construction and Analysis of Systems. pp. 345–357. Springer (1998)
24. Majumdar, R., Saha, I.: Symbolic robustness analysis. In: Real-Time Systems Symposium, 2009, RTSS 2009. 30th IEEE. pp. 355–363. IEEE (2009)
25. Mitra, D.: An asynchronous distributed algorithm for power control in cellular radio systems. In: Wireless and Mobile Communications, pp. 177–186. Springer (1994)
26. Mitra, S., Chandy, K.M.: A formalized theory for verifying stability and convergence of automata in pvs. In: International Conference on Theorem Proving in Higher Order Logics. pp. 230–245. Springer (2008)
27. Olfati-Saber, R., Fax, J.A., Murray, R.M.: Consensus and cooperation in networked multi-agent systems. Proceedings of the IEEE 95(1), 215–233 (2007)
28. Peled, D.: Ten years of partial order reduction. In: International Conference on Computer Aided Verification. pp. 17–28. Springer (1998)
29. Rhee, I.K., Lee, J., Kim, J., Serpedin, E., Wu, Y.C.: Clock synchronization in wireless sensor networks: An overview. Sensors 9(1), 56–85 (2009)
30. Samanta, R., Deshmukh, J.V., Chaudhuri, S.: Robustness analysis of networked systems. In: International Workshop on Verification, Model Checking, and Abstract Interpretation. pp. 229–247. Springer (2013)
31. Welch, J.L., Lynch, N.: A new fault-tolerant algorithm for clock synchronization. Information and computation 77(1), 1–36 (1988)
32. Yang, Y., Chen, X., Gopalakrishnan, G., Kirby, R.M.: Efficient stateful dynamic partial order reduction. In: International SPIN Workshop on Model Checking of Software. pp. 288–305. Springer (2008)