# Symbolic verification of communicating systems with probabilistic message losses: liveness and fairness*

C. Baier[1], N. Bertrand[2], and Ph. Schnoebelen[2]

[1] Universität Bonn, Institut für Informatik I, Germany
[2] LSV, ENS de Cachan & CNRS, France

**Abstract.** NPLCS's are a new model for nondeterministic channel systems where unreliable communication is modeled by probabilistic message losses. We show that, for ω-regular linear-time properties and finite-memory schedulers, qualitative model-checking is decidable. The techniques extend smoothly to questions where fairness restrictions are imposed on the schedulers. The symbolic procedure underlying our decidability proofs has been implemented and used to study a simple protocol handling two-way transfers in an unreliable setting.

## 1 Introduction

*Channel systems* [15] are systems of finite-state components that communicate via asynchronous unbounded fifo channels. *Lossy* channel systems [17, 6], shortly LCS's, are a special class of channel systems where messages can be lost while they are in transit. They are a natural model for fault-tolerant protocols where communication is not supposed to be reliable (see example in Fig. 1 below). Additionally, the lossiness assumption makes termination and safety properties decidable [22, 17, 6, 4, 20, 8] while reliable, i.e., non-lossy, systems are Turing-powerful.

LCS's are a convenient model for verifying safety properties of asynchronous protocols, and this can be automated [4]. However, they are not adequate for verifying liveness and progress properties: firstly these properties are undecidable for LCS's [5], and secondly the model itself is too pessimistic when liveness is considered. Indeed, to ensure any kind of progress, one must assume that at least some messages will not be lost. This is classically obtained via fairness assumptions on message losses [18] but fairness assumptions in LCS's make decidability even more elusive [5, 21].

*Probabilistic LCS's*, shortly PLCS's, are LCS's where message losses are seen as faults having a *probabilistic* behavior [27, 10, 31, 1, 29, 2, 7]. Thanks to its probabilistic framework, this model automatically fulfills strong fairness conditions on the message losses. Additionally it allows one to state so-called *qualitative* questions, whether a linear-time property will be satisfied "*with probability 1*", that are decidable. However, PLCS's are not a realistic model for protocols because they consider that the choices between different actions are made probabilistically rather than nondeterministically. When modeling communication protocols, *nondeterminism* is an essential feature. It

* The first author is supported by the DFG-NWO project VOSS II and the DFG-project PROB-POR. The last two authors were supported by the ACI Sécurité & Informatique project Persée.

is used to model the interleaved behavior of distributed components, to model an unknown environment, to delay implementation choices at early stages of the design, and to abstract away from complex control structures at later stages.

This prompted us to introduce NPLCS's, i.e., channel systems where message losses are probabilistic and actions are nondeterministic [13, 14]. These systems give rise to infinite-state Markov decision processes, and are a more faithful model for analyzing protocols. The drawback is that they raise very difficult verification problems.

**Qualitative verification for NPLCS's.** Our early results in [14] rely on the assumption that idling was always a possible choice. This simplifies the analysis considerably, but is an overkill: a necessary ingredient for most liveness properties of a compound system is the inherent liveness of the components, which disappears if they can idle.

We developed new techniques and removed the idling limitation in [9] where we show that decidability can be maintained if we restrict our attention to *finite-memory* schedulers (strategies for the nondeterministic choices). This seems like a mild restriction, and we adopt it in this paper since we aim for automatic verification.

**Our contributions.** In this paper we extend the preliminary work from [9] in three directions: (1) We allow linear-time formulas referring to the contents of the channels rather than just the control locations. We did not consider this extension earlier because we lacked the techniques for proving the convergence of fixpoint computations. However, the extension is required in practical applications where fairness properties have to express that "a rule is firable," which depends on channel contents for read actions. (2) We develop symbolic representations and algorithms for sets of NPLCS configurations. These algorithms have been implemented in a prototype tool that we use to analyze a simple communication protocol. (3) We consider qualitative verification with quantification over *fair* schedulers, i.e., schedulers that generate fair runs almost surely.

**Outline of the paper.** Section 2 recalls the necessary technical background for nondeterministic probabilistic channel systems, and section 3 introduces the new symbolic framework we use for handling sets of configurations. We present our decidability results in sections 4 (for finite-memory schedulers) and 5 (for fair schedulers). Finally we apply our algorithms on Pachl's protocol in section 6. All proofs omitted in this extended abstract can be found in the complete version available on the web.

## 2 Nondeterministic probabilistic channel systems

We assume the reader has some familiarity with the verification of Markov decision processes, or MDPs, (otherwise see [11]) and refer to [9] for complete definitions regarding our framework. Here we recall the main definitions and notations without motivating or illustrating all of them.

**Lossy channel systems.** A lossy channel system (a LCS) is a tuple $\mathcal{L} = (Q, \mathsf{C}, \mathsf{M}, \Delta)$ of a finite set $Q = \{p, q, \ldots\}$ of control *locations*, a finite set $\mathsf{C} = \{c, \ldots\}$ of *channels*, a finite *message alphabet* $\mathsf{M} = \{m, \ldots\}$ and a finite set $\Delta = \{\delta, \ldots\}$ of *transition rules*. Each rule

has the form $q \xrightarrow{op} p$ where *op* is an *operation* of the form $c!m$ (sending message $m$ along channel $c$), $c?m$ (receiving message $m$ from channel $c$), or $\surd$ (an internal action with no communication). For example, the protocol displayed in Fig 1, is naturally modeled as a LCS: building the asynchronous product of the two processes $P_{\mathsf{L}}$ and $P_{\mathsf{R}}$ yields a bona fide LCS with two channels and a five-message alphabet $\mathsf{M} = \{\mathtt{a}_0, \mathtt{a}_1, \mathtt{d}_0, \mathtt{d}_1, \mathtt{eod}\}$.

**Operational semantics.** A *configuration* of $\mathcal{L}$ as above is a pair $s = (q, w)$ of a location and a channel valuation $w : \mathsf{C} \to \mathsf{M}^*$ associating with any channel its current content (a sequence of messages). $\mathsf{M}^{*\mathsf{C}}$, or $\mathsf{M}^*$ when $|\mathsf{C}| = 1$, denotes the set of all channel valuations, and $\mathsf{Conf}$ the set of all configurations. $\varepsilon$ denotes both the empty word and the empty channel valuation. The size $|s|$ of a configuration is the total number of messages in $s$. The rules of $\mathcal{L}$ give rise to transitions between configurations in the obvious way [9]. We write $\Delta(s)$ for the set of rules $\delta \in \Delta$ that are enabled in configuration $s$.

We write $s \xrightarrow{\delta}_{\text{perf}} s'$ when $s'$ is obtained by firing $\delta$ in $s$. The "perf" subscript stresses the fact that the step is perfect, i.e., no messages are lost. However, in lossy systems, arbitrary messages can be lost. This is formalized with the help of the subword ordering: we write $\mu \sqsubseteq \mu'$ when $\mu$ is a subword of $\mu'$, i.e., $\mu$ can be obtained by removing (any number of) messages from $\mu'$, and we extend this to configurations, writing $(q, w) \sqsubseteq (q', w')$ when $q = q'$ and $w(c) \sqsubseteq w'(c)$ for all $c \in \mathsf{C}$. As a consequence of Higman's Lemma, $\sqsubseteq$ is a well-quasi-order (a *wqo*) between configurations of $\mathcal{L}$. Now, we define *lossy steps* by letting $s \xrightarrow{\delta} s'' \stackrel{\text{def}}{\Leftrightarrow}$ there is a perfect step $s \xrightarrow{\delta}_{\text{perf}} s'$ such that $s'' \sqsubseteq s'$. This gives rise to a labeled transition system $LTS_{\mathcal{L}} \stackrel{\text{def}}{=} (\mathsf{Conf}, \Delta, \to)$. Here the set $\Delta$ of transition rules serves as action alphabet. In the following we assume that for any location $q \in Q$, $\Delta$ contains at least one rule $q \xrightarrow{op} p$ where *op* is not a receive operation. This hypothesis ensures that $LTS_{\mathcal{L}}$ has no deadlock configuration and makes the theory smoother. It is no real loss of generality as demonstrated in [2, § 8.3].

**An example.** Pachl's protocol [22] handles two-way communications over lossy channels and is our case study for our algorithms. It consists of two identical processes, $P_{\mathtt{L(eft)}}$ and $P_{\mathtt{R(ight)}}$, that exchange data over lossy channels using an acknowledgment mechanism based on the alternating bit protocol. See Fig 1 below. The actual contents of the data messages is abstracted away, and we just use $\mathtt{d}_0, \mathtt{d}_1 \in \mathsf{M}$ to record the alternating control bit. Messages $\mathtt{a}_0, \mathtt{a}_1 \in \mathsf{M}$ are the corresponding acknowledgments. The protocol starts in configuration $(\mathtt{L0}, \mathtt{R4})$ where $P_{\mathsf{L}}$ is the sender and $P_{\mathsf{R}}$ the receiver. At any time (provided its last data message has been acknowledged) the sender may signal the end of its data sequence with the $\mathtt{eod} \in \mathsf{M}$ control message and then the two processes swap their sending and receiving roles. Note that $\mathtt{eod}$ does not need to carry a control bit, and that its correct reception is not acknowledged. In section 6 we explain how such a two-process protocol is modeled as an LCS, and give some outcomes of our automated analysis.

**From LCS's to NPLCS's.** A NPLCS $\mathcal{N} = (\mathcal{L}, \tau)$ is a LCS $\mathcal{L}$ further equipped with a *fault rate* $\tau \in (0, 1)$ that specifies the probability that a given message stored in one of the message queues is lost during a step [13, 14]. The operational semantics of NPLCS's

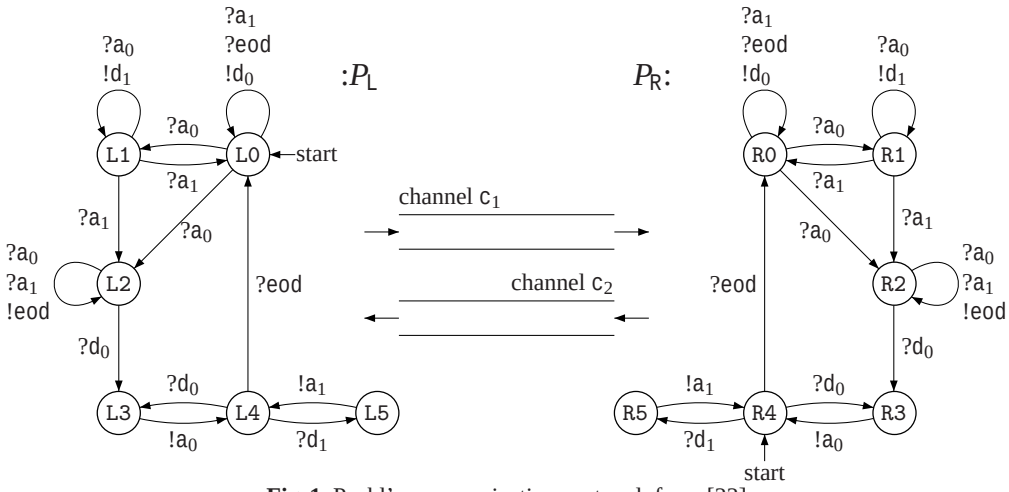


**Fig. 1.** Pachl's communication protocol, from [22].

has the form of an infinite-state Markov decision process $MDP_{\mathcal{N}} \stackrel{\text{def}}{=} (\mathsf{Conf}, \Delta, \mathbf{P}_{\mathcal{N}})$. The stepwise probabilistic behavior is formalized by a three-dimensional transition probability matrix $\mathbf{P}_{\mathcal{N}} : \mathsf{Conf} \times \Delta \times \mathsf{Conf} \to [0,1]$. For a given configuration $s$ and an enabled rule $\delta \in \Delta(s)$, $\mathbf{P}_{\mathcal{N}}(s,\delta,\cdot)$ is a distribution over $\mathsf{Conf}$, while $\mathbf{P}_{\mathcal{N}}(s,\delta,\cdot) = 0$ for any transition rule $\delta$ that is not enabled in $s$. The intuitive meaning of $\mathbf{P}_{\mathcal{N}}(s,\delta,t) = \lambda > 0$ is that with probability $\lambda$, the system moves from configuration $s$ to configuration $t$ when $\delta$ is the chosen transition rule in $s$.

For lack of space, this extended abstract omits the technically heavy but quite natural definition of $\mathbf{P}_{\mathcal{N}}$, and only lists its two essential properties:
1. the labeled transition system underlying $MDP_{(\mathcal{L},\tau)}$ is exactly $LTS_{\mathcal{L}}$.
2. the set $Q_\varepsilon = \{(q,\varepsilon) \mid q \in Q\}$ of configurations where the channels are empty is an *attractor*, i.e., from any starting configuration, $Q_\varepsilon$ will eventually be visited with probability 1 [2, 7].

**Schedulers and probability measure.** The nondeterminism in an MDP is resolved by a *scheduler*, also often called "adversary", "policy" or "strategy". Here a "scheduler" is a *history-dependent deterministic scheduler* in the classification of [28]. Formally, a scheduler for $\mathcal{N}$ is a mapping $\mathcal{U}$ that assigns to any finite path $\pi$ in $\mathcal{N}$ a transition rule $\delta \in \Delta$ that is enabled in the last state of $\pi$. The given path $\pi$ specifies the history of the system, and $\mathcal{U}(\pi)$ is the rule that $\mathcal{U}$ chooses to fire next. A scheduler $\mathcal{U}$ only gives rise to certain paths: we say $\pi = s_0 \xrightarrow{\delta_1} s_1 \xrightarrow{\delta_2} \cdots$ is *compatible with* $\mathcal{U}$ or, shortly, is a $\mathcal{U}$*-path*, if $\mathbf{P}_{\mathcal{N}}(s_{i-1},\delta_i,s_i) > 0$ for all $i \geq 1$, where $\delta_{i+1} = \mathcal{U}(s_0 \xrightarrow{\delta_1} \cdots \xrightarrow{\delta_i} s_i)$ is the rule chosen by $\mathcal{U}$ at step $i$ along $\pi$. In practice, it is only relevant to define how $\mathcal{U}$ evaluates on finite $\mathcal{U}$-paths.

A *finite-memory*, or fm-, scheduler $\mathcal{U} = (U, D, \eta, u_0)$ is specified via a finite set $U$ of *modes*, a *starting mode* $u_0 \in U$, a *decision rule* $D : U \times \mathsf{Conf} \to \Delta$ choosing the next rule $D(u,s) \in \Delta(s)$ based on the current mode and the current configuration, and a *next-mode function* $\eta : U \times \mathsf{Conf} \to U$ specifying the mode-changes of $\mathcal{U}$. The modes are used to

store some relevant information about the history. An fm-scheduler $\mathcal{U}$ is *memoryless* if it has a single mode: then $\mathcal{U}$ is not history-dependent and can be specified more simply as a mapping $\mathcal{U} : \mathsf{Conf} \to \Delta$.

Now, given an NPLCS $\mathcal{N}$, a starting configuration $s = s_0$ and a scheduler $\mathcal{U}$, the behavior of $\mathcal{N}$ under $\mathcal{U}$ can be formalized by an infinite-state Markov chain $MC_{\mathcal{U}}$. For arbitrary schedulers, the states of $MC_{\mathcal{U}}$ are finite paths in $\mathcal{N}$, while for fm-schedulers it is possible to consider pairs $(u,s)$ of a mode of $\mathcal{U}$ and a configuration of $\mathcal{N}$. One may now apply the standard machinery for Markov chains and define (for fixed starting configuration $s$) a sigma-field on the set of infinite paths starting in $s$ and a probability measure on it, see, e.g., [28, 23, 11]. We shall write $\Pr_{\mathcal{U}}\left(s \models \ldots\right)$ to denote the standard probability measure in $MC_{\mathcal{U}}$ with starting state $s$.

**LTL/CTL-notation.** We use simple LTL and CTL formulas to denote properties of respectively paths and configurations in $MDP_{\mathcal{L}}$. Here configurations and locations serve as atomic propositions: for example $\Box\Diamond s$ (resp. $\Box\Diamond q$) means that $s \in \mathsf{Conf}$ (resp. $q \in Q$) is visited infinitely many times, and $q$ Until $s$ means that the control location remains $q$ until configuration $s$ is eventually reached. These notations extend to sets and, for $T \subseteq \mathsf{Conf}$ and $P \subseteq Q$, $\Box\Diamond T$ and $\Box\Diamond P$ have the obvious meaning. For $P \subseteq Q$, $P_\varepsilon$ is the set $\{(p,\varepsilon) \mid p \in P\}$ so that $\Diamond Q_\varepsilon$ means that eventually a configuration with empty channels is reached. It is well-known that for any scheduler $\mathcal{U}$, the set of paths starting in some configuration $s$ and satisfying an LTL formula, or an $\omega$-regular property, $\varphi$ is measurable [32, 16]. We write $\Pr_{\mathcal{U}}\left(s \models \varphi\right)$ for this measure.

**Reachability analysis.** For a set $A \subseteq \mathsf{Conf}$ and a rule $\delta \in \Delta$, we let $\mathit{Pre}[\delta](A) \stackrel{\text{def}}{=} \{s \mid \exists t \in A, s \xrightarrow{\delta} t\}$ denote the set of configurations from which $A$ can be reached in one step with rule $\delta$. $\mathit{Pre}(A) \stackrel{\text{def}}{=} \bigcup_{\delta\in\Delta} \mathit{Pre}[\delta](A)$ contains all *one-step predecessors*, and $\mathit{Pre}^*(A) \stackrel{\text{def}}{=} A \cup \mathit{Pre}(A) \cup \mathit{Pre}(\mathit{Pre}(A)) \cup \cdots$ all *iterated predecessors*. The successor sets $\mathit{Post}[\delta](A)$, $\mathit{Post}(A)$, and $\mathit{Post}^*(A)$ are defined analogously. Recall that reachability between configurations of LCS's is decidable [6, 30], which is also implied by Theorem 3.2 below.

**Constrained reachability.** We sometimes need to reach a set $A$ using only rules that cannot get us out of some set $T \subseteq \mathsf{Conf}$. Formally, for $T, A \subseteq \mathsf{Conf}$, we define

$$\widehat{\mathit{Pre}}_T(A) \stackrel{\text{def}}{=} \{s \in \mathsf{Conf} \mid \exists \delta \in \Delta(s) \text{ s.t. } \mathit{Post}[\delta](s) \cap A \neq \emptyset \text{ and } \mathit{Post}[\delta](s) \subseteq T\}.$$

In other words, $s$ is in $\widehat{\mathit{Pre}}_T(A)$ if there is a rule $\delta$ that may take $s$ to some state in $A$ but that cannot take it outside $T$. The set of iterated *$T$-constrained* predecessors is $\widehat{\mathit{Pre}}^*_T(A) \stackrel{\text{def}}{=} A \cup \widehat{\mathit{Pre}}_T(A) \cup \widehat{\mathit{Pre}}_T(\widehat{\mathit{Pre}}_T(A)) \cup \cdots$

## 3 Symbolic representations for sets of configurations

Symbolic model-checking relies on symbolic objects representing sets of configurations, and algorithmic methods for handling these objects meaningfully.

In this section, we present a symbolic framework for NPLCS's based on *differences of prefixed upward-closures*. This extends previous techniques from [4, 3, 20] in that it

permits dealing with set differences and checking which is the first message in a channel. For simplicity in the presentation, *we assume that the NPLCS under consideration only has a single channel*. We also omit most of the algorithmic details pertaining to data structures, normal forms, canonization, ..., that are present in our prototype implementation (see section 6).

Recall that a set $T \subseteq \mathsf{Conf}$ is *upward-closed* (resp., *downward-closed*) if for all $s \in T$, and for all $s' \sqsupseteq s$ (resp., $s' \sqsubseteq s$), $s' \in T$. For $T \subseteq \mathsf{Conf}$, we let $\uparrow T \stackrel{\text{def}}{=} \{s \in \mathsf{Conf} | \exists s' \in T \wedge s' \sqsubseteq s\}$ denote the *upward-closure* of $T$, and $\downarrow T \stackrel{\text{def}}{=} \{s \in \mathsf{Conf} | \exists s' \in T \wedge s \sqsubseteq s'\}$ denote the *downward-closure* of $T$. For singleton sets we write shortly $\uparrow t$ and $\downarrow t$ rather than $\uparrow \{t\}$ and $\downarrow \{t\}$.

Our symbolic sets are defined with the following abstract grammar:

| | | |
|---|---|---|
| prefix: | $\alpha := \varepsilon \mid m$ | $m \in \mathsf{M}$ |
| prefixed closure: | $\theta := \alpha{\uparrow}u$ | $u \in \mathsf{M}^*$ |
| sum of prefixed closures: | $\sigma := \theta_1 + \cdots + \theta_n$ | $n \geq 0$ |
| simple symbolic set: | $\rho := \langle q, \theta - \sigma \rangle$ | $q \in Q$ is a location |
| symbolic set: | $\gamma := \rho_1 + \cdots + \rho_n$ | $n \geq 0$ |

Prefixed (upward-)closures and their sums denote subsets of $\mathsf{M}^*$ defined with $[\![\alpha{\uparrow}u]\!] \stackrel{\text{def}}{=} \{\alpha v \mid u \sqsubseteq v\}$ and $[\![\theta_1 + \cdots + \theta_n]\!] \stackrel{\text{def}}{=} [\![\theta_1]\!] \cup \cdots \cup [\![\theta_n]\!]$. Symbolic sets denote subsets of $\mathsf{Conf}$ defined with $[\![\langle q, \theta - (\theta_1 + \cdots + \theta_n)\rangle]\!] \stackrel{\text{def}}{=} \{\langle q, v\rangle \in \mathsf{Conf} \mid v \in [\![\theta]\!] \smallsetminus ([\![\theta_1]\!] \cup \cdots \cup [\![\theta_n]\!])\}$. A *region* is any subset of $\mathsf{Conf}$ that can be denoted by a symbolic set. It is a *control region* if can be written under the form $\sum_i \langle q_i, \varepsilon{\uparrow}\varepsilon\rangle$, where channel contents are unrestricted.

We abuse notation and write $\emptyset$ to denote both empty (i.e., with $n = 0$) sums of prefixed closures and empty symbolic sets. We also sometimes write $\uparrow v$ for $\varepsilon{\uparrow}v$, $\theta - \theta_1 - \cdots - \theta_n$ for $\theta - (\theta_1 + \cdots + \theta_n)$, and $\theta$ for $\theta - \emptyset$. We write $\gamma \equiv \gamma'$ when $[\![\gamma]\!] = [\![\gamma']\!]$, i.e., when $\gamma$ and $\gamma'$ denote the same region.

**Theorem 3.1 (Effective symbolic computation: basics).**

**Boolean closure:** *Regions are closed under union, intersection, and complementation. Moreover, there exist algorithms that given symbolic sets $\gamma_1$ and $\gamma_2$ return terms denoted $\gamma_1 \sqcup \gamma_2$, $\gamma_1 \sqcap \gamma_2$ and $\neg\gamma$ such that $[\![\gamma_1 \sqcup \gamma_2]\!] = [\![\gamma_1]\!] \cup [\![\gamma_2]\!]$, $[\![\gamma_1 \sqcap \gamma_2]\!] = [\![\gamma_1]\!] \cap [\![\gamma_2]\!]$ and $[\![\neg\gamma]\!] = \mathsf{Conf} \smallsetminus [\![\gamma]\!]$.*

**Upward closure:** *Regions are closed under upward closure. Moreover, there exists an algorithm that given a symbolic set $\gamma$ returns a term denoted $\uparrow\gamma$ such that $[\![\uparrow\gamma]\!] = \uparrow[\![\gamma]\!]$.*

**Vacuity:** *It is decidable whether $[\![\gamma]\!] = \emptyset$ given a region $\gamma$.*

**One-step predecessors:** *Regions are closed under the $Pre(\_)$ and $\widehat{Pre}_{\_}(\_)$ operations. Moreover, there exist algorithms that given symbolic sets $\gamma$ and $\gamma'$ return terms denoted $Pre(\gamma)$ and $\widehat{Pre}_{\gamma'}(\gamma)$, and such that $[\![Pre(\gamma)]\!] = Pre([\![\gamma]\!])$ and $[\![\widehat{Pre}_{\gamma'}(\gamma)]\!] = \widehat{Pre}_{[\![\gamma']\!]}([\![\gamma]\!])$.*

Theorem 3.1 provides the basic ingredients necessary for symbolic model-checking of LCS's. These ingredients can then be used for computing sets defined as fixpoints.

For example, using standard $\mu$-calculus notation, a symbolic set denoting $Pre^*(\llbracket\gamma\rrbracket)$ would be defined as $\mu X.\gamma \sqcup Pre(X)$. In [8] we show how a symbolic representation for sets defined by such fixpoint expressions can be computed effectively (when some guardedness condition holds).

**Theorem 3.2 (Effective symbolic computation: fixpoints).**

**Iterated (constrained) predecessors:** *Regions are closed under the $Pre^*(\_)$ and the $\widehat{Pre}^*_\_(\_)$ operations. Moreover, there exist algorithms that given symbolic sets $\gamma$ and $\gamma'$ return terms denoted $Pre^*(\gamma)$ and $\widehat{Pre}^*_{\gamma'}(\gamma)$, and such that $\llbracket Pre^*(\gamma)\rrbracket = Pre^*(\llbracket\gamma\rrbracket)$ and $\llbracket\widehat{Pre}^*_{\gamma'}(\gamma)\rrbracket = \widehat{Pre}^*_{\llbracket\gamma'\rrbracket}(\llbracket\gamma\rrbracket)$.*

**Safe sets (see section 4):** *For any region $\gamma$, the set $\nu X.\big(\gamma \sqcap \widehat{Pre}_X(\mathsf{Conf})\big)$ is a region, and a term for it can be computed effectively.*

**Promising sets (see section 4):** *For any region $\gamma$, the set $\nu X.\widehat{Pre}^*_X(\gamma)$ is a region, and a term for it can be computed effectively.*

**$\exists$CTL:** *The set of configurations satisfying an $\exists$CTL formula (i.e., a CTL formula where only the modalities "$\exists(\_$ Until $\_)$" and "$\exists$Next $\_$" are allowed) is a region when the atomic propositions are themselves regions. Moreover, a symbolic set for that region can be obtained algorithmically from the $\exists$CTL formula.*

## 4 Verifying safety and liveness properties for NPLCS's

This section considers various types of safety and liveness properties where regions serve as atoms, and presents algorithms for checking the existence of a fm-scheduler $\mathcal{U}$ such that $\Pr_{\mathcal{U}}(s \models \varphi)$ is $> 0$, $= 1$, $< 1$ or $= 0$.

We start with reachability properties $\Diamond A$ and invariants $\Box A$ for some region $A$.

For eventually properties with the satisfaction criteria "with positive probability", decidability relies on the computation of iterative predecessors in (non-probabilistic) lossy channel systems:

**Theorem 4.1.** *Let $s \in \mathsf{Conf}$ and $A \subseteq \mathsf{Conf}$. There exists a scheduler $\mathcal{U}$ with $\Pr_{\mathcal{U}}(s \models \Diamond A) > 0$ iff $\Pr_{\mathcal{U}}(s \models \Diamond A) > 0$ for some memoryless scheduler $\mathcal{U}$ iff $s \in Pre^*(A)$.*

For other satisfaction criteria, or for other properties, we have to develop more ad-hoc characterizations of the sets of configurations where the qualitative properties hold.

For invariants $\Box A$, we introduce the concept of safe sets:

**Definition 4.2 (Safe sets).** *Let $A, T \subseteq \mathsf{Conf}$. $T$ is called* safe *for $A$ if $T \subseteq A$ and for all $s \in T$, there exists a transition rule $\delta$ enabled in $s$ such that $Post[\delta](s) \subseteq T$.*

Since the union of safe sets is safe, the largest safe set for $A$, denoted $Safe(A)$, exists.

There exists a simple fixpoint characterization for $Safe(A)$ (here and in the sequel, we use the standard $\mu/\nu$-notations for fixpoints).

**Lemma 4.3.** *For any $A \subseteq \mathsf{Conf}$, $Safe(A) = \nu X.A \cap \widehat{Pre}_X(\mathsf{Conf})$.*

Thus, if $A$ is a region, $Safe(A)$ is a region too, and a symbolic representation can be computed effectively (Theorem 3.2). This is the key for verifying invariants:

**Theorem 4.4 (Safe sets and invariants).** *Let $A \subseteq \mathsf{Conf}$ and $s \in \mathsf{Conf}$.*

*(a)* $s \in \mathit{Safe}(A)$
iff *there exists a scheduler $\mathcal{U}$ such that* $\Pr_{\mathcal{U}}(s \models \Box A) = 1$
iff *there exists a memoryless scheduler $\mathcal{U}$ such that* $\Pr_{\mathcal{U}}(s \models \Box A) = 1$.
*(b)* $s \models \exists(A \text{ Until } \mathit{Safe}(A))$
iff *there exists a scheduler $\mathcal{U}$ such that* $\Pr_{\mathcal{U}}(s \models \Box A) > 0$
iff *there exists a memoryless scheduler $\mathcal{U}$ such that* $\Pr_{\mathcal{U}}(s \models \Box A) > 0$.

The corollary is that, for a region $A$, we can compute a symbolic representation for the set of all configurations where $\Pr_{\mathcal{U}}(s \models \Box A) > 0$ or $= 1$ for some scheduler $\mathcal{U}$.

**Definition 4.5 (Promising sets).** *Let $A, T \subseteq \mathsf{Conf}$. $T$ is called* promising *for $A$ if for all $s \in T$ there exists a path $s = s_0 \xrightarrow{\delta_1} s_1 \xrightarrow{\delta_2} \cdots \xrightarrow{\delta_m} s_m$ with $m \geq 0$ such that $s_m \in A$ and for all $1 \leq i \leq m$, $\mathit{Post}[\delta_i](s_{i-1}) \subseteq T$.*

As for safe sets, the largest promising set for $A$ exists: we denote it $\mathit{Prom}(A)$.

**Lemma 4.6.** *For any $A \subseteq \mathsf{Conf}$, $\mathit{Prom}(A) = \nu X.\widehat{Pre}^*_X(A)$.*

Thus, if $A$ is a region, $\mathit{Prom}(A)$ is a region too, and a symbolic representation can be computed effectively (Theorem 3.2).

**Theorem 4.7 (Promising sets and almost sure reachability).** *Let $s \in \mathsf{Conf}$ and $A \subseteq \mathsf{Conf}$. $s \in \mathit{Prom}(A)$ iff $\Pr_{\mathcal{U}}(s \models \Diamond A) = 1$ for some scheduler $\mathcal{U}$ iff $\Pr_{\mathcal{U}}(s \models \Diamond A) = 1$ for some memoryless scheduler $\mathcal{U}$.*

The corollary is that, for a region $A$, we can compute the set of all configurations $s$ such that $\Pr_{\mathcal{U}}(s \models \Diamond A) > 0$ or $= 1$ for some $\mathcal{U}$.

We now consider repeated reachability and persistence properties. The question whether a repeated reachability property $\Box\Diamond A$ holds under some scheduler with positive probability is undecidable when ranging over the full class of schedulers, but is decidable for the class of fm-schedulers. This was shown in [14, 9] for the case where $A$ is a set of locations (*i.e.* a control region). We now show that the decidability even holds if $A$ is a region. More precisely, we show that if $A$ is a region and $\varphi \in \{\Box\Diamond A, \Diamond\Box A\}$, then the set of configurations $s$ where $\Pr_{\mathcal{U}}(s \models \varphi) > 0$ or $= 1$ for some fm-scheduler is a region.

For $A \subseteq \mathsf{Conf}$ let $\mathit{Prom}^{\geq 1}(A)$ denote the largest set $T$ of configurations such that for all $t \in T$ there exists a finite path $s = s_0 \xrightarrow{\delta_1} s_1 \xrightarrow{\delta_2} \cdots \xrightarrow{\delta_m} s_m$ with $m \geq 1$, $s_m \in A$ and $\mathit{Post}[\delta_i](s_{i-1}) \subseteq T$ for all $1 \leq i \leq m$. Note that the definition of $\mathit{Prom}^{\geq 1}(A)$ is different from $\mathit{Prom}(A)$ since the paths must have length at least 1. We then have $\mathit{Prom}^{\geq 1}(A) = \nu X.\widehat{Pre}^+_X(A)$, and, if $A$ is a region then so is $\mathit{Prom}^{\geq 1}(A)$. Thus, the following theorem provides the decidability of repeated reachability and persistence properties:

**Theorem 4.8 (Repeated reachability and persistence).** *Let $s \in \mathsf{Conf}$ and $A \subseteq \mathsf{Conf}$.*

*(a)* $s \in \mathit{Prom}^{\geq 1}(A)$ *iff* $\Pr_{\mathcal{U}}(s \models \Box\Diamond A) = 1$ *for some scheduler* $\mathcal{U}$
*iff* $\Pr_{\mathcal{U}}(s \models \Box\Diamond A) = 1$ *for some memoryless scheduler* $\mathcal{U}$.
*(b)* $s \in \mathit{Pre}^*(\mathit{Prom}^{\geq 1}(A))$ *iff* $\Pr_{\mathcal{U}}(s \models \Box\Diamond A) > 0$ *for some fm-scheduler* $\mathcal{U}$
*iff* $\Pr_{\mathcal{U}}(s \models \Box\Diamond A) > 0$ *for some memoryless scheduler* $\mathcal{U}$.
*(c)* $s \in \mathit{Prom}(\mathit{Safe}(A))$ *iff* $\Pr_{\mathcal{U}}(s \models \Diamond\Box A) = 1$ *for some scheduler* $\mathcal{U}$
*iff* $\Pr_{\mathcal{U}}(s \models \Diamond\Box A) = 1$ *for some memoryless scheduler* $\mathcal{U}$.
*(d)* $s \in \mathit{Pre}^*(\mathit{Safe}(A))$ *iff* $\Pr_{\mathcal{U}}(s \models \Diamond\Box A) > 0$ *for some scheduler* $\mathcal{U}$
*iff* $\Pr_{\mathcal{U}}(s \models \Diamond\Box A) > 0$ *for some memoryless scheduler* $\mathcal{U}$.

We now consider the Streett formula $\varphi_S = \bigwedge_{1\leq i\leq n} \Box\Diamond A_i \rightarrow \Box\Diamond B_i$ where $A_1,\ldots,A_n$ and $B_1,\ldots,B_n$ are regions. Here again we only consider fm-schedulers since the problem is undecidable for the full class of schedulers [9].

For $A, B \subseteq \mathsf{Conf}$, let $\mathit{Prom}_A^{\geq 1}(B)$ be the largest subset $T$ of $A$ such that for all $t \in T$ there exists a path $t = s_0 \xrightarrow{\delta_1} \cdots \xrightarrow{\delta_m} s_m$ with $m > 0$, $s_m \in B$ and $\mathit{Post}[\delta_i](s_{i-1}) \subseteq T$ for all $1 \leq i \leq m$. We have $\mathit{Prom}_A^{\geq 1}(B) = \nu X.\widehat{\mathit{Pre}}_X^+(B) \cap A$ and if $A, B$ are regions then so is $\mathit{Prom}_A^{\geq 1}(B)$. In addition, $s \in \mathit{Prom}_A^{\geq 1}(B)$ iff $\Pr_{\mathcal{U}}(s \models \Box\Diamond B \wedge \Box A) = 1$ for some fm-scheduler $\mathcal{U}$.

The above is useful to show decidability of the questions whether $\Pr_{\mathcal{U}}(s \models \varphi_S) < 1$ or $= 0$ for some fm-scheduler $\mathcal{U}$. For this, we use the fact that $\Pr_{\mathcal{U}}(s \models \varphi_S) < 1$ iff $\Pr_{\mathcal{U}}(s \models \Box\Diamond A_i \rightarrow \Box\Diamond B_i) < 1$ for some $i$ iff $\Pr_{\mathcal{U}}(s \models \Box\Diamond A_i \wedge \Diamond\Box\neg B_i) > 0$ for some $i$.

**Theorem 4.9 (Streett property, probability less than 1).** *There exists a fm-scheduler* $\mathcal{U}$ *with* $\Pr_{\mathcal{U}}(s \models \Box\Diamond A \wedge \Diamond\Box\neg B) > 0$ *iff there exists a memoryless scheduler* $\mathcal{U}$ *with* $\Pr_{\mathcal{U}}(s \models \Box\Diamond A \wedge \Diamond\Box\neg B) > 0$ *iff* $s \in \mathit{Pre}^*(\mathit{Prom}_{\neg B}^{\geq 1}(A))$. *In particular,* $\Pr_{\mathcal{U}}(s \models \varphi_S) < 1$ *for some fm-scheduler* $\mathcal{U}$ *iff* $s \in \bigcup_{1\leq i\leq n} \mathit{Pre}^*(\mathit{Prom}_{\neg B_i}^{\geq 1}(A_i))$.

Let $T_i$ be the set of all configurations $t \in \mathsf{Conf}$ such that $\Pr_{\mathcal{W}}(s \models \Box\Diamond A_i \wedge \Diamond\Box\neg B_i) = 1$ for some fm-scheduler $\mathcal{W}$. Note that $T_i = \mathit{Pre}^*(\mathit{Prom}_{\neg B_i}^{\geq 1}(A_i))$ is a region. Thus, $T_S = T_1 \cup T_2 \cup \cdots \cup T_n$ is a region too. This and the following theorem yields the decidability of the question whether $\Pr_{\mathcal{U}}(s \models \varphi_S) = 0$ for some scheduler $\mathcal{U}$.

**Theorem 4.10 (Streett property, zero probability).** *There exists a fm-scheduler* $\mathcal{U}$ *such that* $\Pr_{\mathcal{U}}(s \models \varphi_S) = 0$ *if and only if* $s \in \mathit{Prom}(T_S)$.

We next consider the satisfaction criterion "with positive probability". The treatment of the special case of a single strong fairness formula $\Box\Diamond A \rightarrow \Box\Diamond B \equiv \Diamond\Box\neg A \vee \Box\Diamond B$ is obvious as we have: There exists a finite-memory (resp. memoryless) scheduler $\mathcal{U}$ such that $\Pr_{\mathcal{U}}(s \models \Box\Diamond A \rightarrow \Box\Diamond B) > 0$ iff at least one of the following conditions holds: (i) there exists a fm-scheduler $V$ such that $\Pr_{\mathcal{V}}(s \models \Diamond\Box\neg A) > 0$ or (ii) there exists a fm-scheduler $\mathcal{W}$ such that $\Pr_{\mathcal{W}}(s \models \Box\Diamond B) > 0$. We now extend this observation to the general case (several Streett properties). For $I \subseteq \{1,\ldots,n\}$, let $A_I$ denote the set of configurations $s$ such that there exists a finite-memory scheduler satisfying $\Pr_{\mathcal{U}}(s \models \bigwedge_{i\in I} \Box\Diamond B_i \wedge \bigwedge_{i\notin I} \Box\neg A_i) = 1$ and let $A$ be the union of all $A_I$'s, i.e., $A = \bigcup_{I\subseteq\{1,\ldots,n\}} A_I$. Then, the sets $A_I$ and $A$ are regions. Thus, the algorithmic treatment of Streett properties the satisfaction criteria "positive probability" and "almost surely" relies on the following theorem:

**Theorem 4.11 (Streett properties, positive probability and almost surely).**

*(a) There exists a fm-scheduler $\mathcal{U}$ such that $\Pr_{\mathcal{U}}(s \models \varphi_S) > 0$ iff $s \in Pre^*(A)$.*
*(b) There exists a fm-scheduler $\mathcal{U}$ such that $\Pr_{\mathcal{U}}(s \models \varphi_S) = 1$ iff $s \in Prom(A)$.*

We conclude with the following main theorem gathering all previous results:

**Theorem 4.12 (Qualitative model-checking).** *For any NPLCS $\mathcal{N}$ and Streett property $\varphi = \bigwedge_i \Box\Diamond A_i \to \Box\Diamond B_i$ where the $A_i$'s and $B_i$'s are regions, the set of all configurations $s$ s.t. for all fm-schedulers $\mathcal{U}$ $\Pr_{\mathcal{U}}(s \models \varphi)$ satisfies a qualitative constraint "$=1$", or "$<1$", or "$=0$", or "$>0$", is a region that can be computed effectively.*

With the techniques of [9, § 7], Theorem 4.12 extends to all ω-regulars properties

## 5 Verification under fair finite-memory schedulers

We now address the problem of verifying qualitative linear time properties under fairness assumptions. Following the approaches of [19, 32, 12], we consider here a notion of *scheduler-fairness* which rules out some schedulers that generate unfair paths with positive probability. This notion of scheduler-fairness has to be contrasted with extreme- and alpha-fairness introduced in [24–26] which require a "fair" resolution of probabilistic choices and serve as verification techniques rather than fairness assumptions about the nondeterministic choices.

A scheduler $\mathcal{U}$ is called *fair* if it generates almost surely fair paths, according to some appropriate fairness constraints for paths. We deal here with *strong fairness* for selected sets of transition rules. I.e., we assume a set $\mathcal{F} = \{f_0, \ldots, f_{k-1}\}$ where $f_i \subseteq \Delta$ and require strong fairness for all $f_i$'s. (The latter means whenever some transition rule in $f_i$ is enabled infinitely often then some transition rule in $f_i$ will fire infinitely often.) For instance, process fairness for $k$ processes $P_0, \ldots, P_{k-1}$ can be modelled by $\mathcal{F} = \{f_0, \ldots, f_{k-1}\}$ where $f_i$ is the set of transition rules describing $P_i$'s actions.

A set $f \subseteq \Delta$ is called enabled in configuration $s$ if there is a transition rule $\delta \in f$ that is enabled in $s$, i.e., if $\Delta(s) \cap f \neq \emptyset$. If $F$ is a subset of $\mathcal{F}$ and $s \in \mathsf{Conf}$ then $F$ is called enabled in $s$ if some $f \in F$ is enabled in $s$, i.e., if $\exists f \in F.\, f \cap \Delta(s) \neq \emptyset$. We write $\mathsf{Enabl}(F)$ to denote the set of configurations $s \in \mathsf{Conf}$ where $F$ is enabled.

**Definition 5.1 (Fair paths, fair schedulers).** *Let $\mathcal{F} \in 2^{2^\Delta}$ be a (finite) set consisting of subsets of $\Delta$. An infinite path $s_0 \xrightarrow{\delta_1} s_1 \xrightarrow{\delta_2} \cdots$ is called* $\mathcal{F}$-fair *iff for all $f \in \mathcal{F}$ either $\delta_j \in f$ for infinitely many $j$ or there is some $i \geq 0$ such that $f$ is not enabled in the configurations $s_j$ for all $j \geq i$. Scheduler $\mathcal{U}$ is called* $\mathcal{F}$-fair *(or briefly* fair*) if for each starting state $s$, almost all $\mathcal{U}$-paths are $\mathcal{F}$-fair.*

We first consider *reachability* properties $\Diamond A$ and show that fairness assumptions are irrelevant for the satisfaction criteria "with positive probability"and "almost surely". This follows from the fact that from the moment on where a configuration in $A$ has been entered one can continue in an arbitrary, but $\mathcal{F}$-fair way. Thus:

$$\exists\, \mathcal{V}\ \mathcal{F}\text{-fair s.t. } \Pr_{\mathcal{V}}(s \models \Diamond A) > 0 \quad \text{iff} \quad \exists\, \mathcal{U} \text{ s.t. } \Pr_{\mathcal{U}}(s \models \Diamond A) > 0$$
$$\exists\, \mathcal{V}\ \mathcal{F}\text{-fair s.t. } \Pr_{\mathcal{V}}(s \models \Diamond A) = 1 \quad \text{iff} \quad \exists\, \mathcal{U} \text{ s.t. } \Pr_{\mathcal{U}}(s \models \Diamond A) = 1$$

By the results of section 4, given an NPLCS $\mathcal{N}$, starting configuration $s$ and region $A$, the questions whether there exists a $\mathcal{F}$-fair scheduler $\mathcal{U}$ such that $\Pr_{\mathcal{U}}(s \models \Diamond A) > 0$ or $= 1$ are decidable.

The treatment of *invariant* properties $\Box A$ under fairness constraints relies on generalizations of the concept of safe and promising sets. For $A, B \subseteq \mathsf{Conf}$, $\mathit{Prom}_A(B)$ denotes the largest set $T \subseteq A \cup B$ such that for all $t \in T$ there exists a path $t = s_0 \xrightarrow{\delta_1} \cdots \xrightarrow{\delta_m} s_m$ with $m \geq 0$, $s_m \in B$ and $\mathit{Post}[\delta_i](s_{i-1}) \subseteq T$ for all $1 \leq i \leq m$. The fixed-point definition of $\mathit{Prom}_A(B)$ would be $\nu X.\widehat{Pre}_X^*(B) \cap (A \cup B)$.

For $\mathcal{F} \subseteq 2^{\Delta}$ and $A \subseteq \mathsf{Conf}$, let $\mathit{Safe}_{\mathcal{F}}(A) = \bigcup_{F \subseteq \mathcal{F}} \mathit{Safe}[F](A)$ where $\mathit{Safe}[F](A)$ is defined as follows. If $F$ is a nonempty subset of $\mathcal{F}$ then $\mathit{Safe}[F](A)$ denotes the largest set $T \subseteq A \setminus \mathsf{Enabl}(\mathcal{F} \setminus F)$ such that for all $t \in T$ and $f \in F$ there is a path $s_0 \xrightarrow{\delta_1} \cdots \xrightarrow{\delta_m} s_m$ with $t = s_0$, $m \geq 1$, $\delta_m \in f$ and $\mathit{Post}[\delta_i](s_{i-1}) \subseteq T$ for all $1 \leq i \leq m$. Moreover, $\mathit{Safe}_{\mathcal{F}}[\emptyset](A) = \mathit{Safe}(A \setminus \mathsf{Enabl}(\mathcal{F}))$.

Since $\mathsf{Enabl}(\mathcal{F} \setminus F)$ can be expressed by $\mathit{Pre}[\mathcal{F} \setminus F](\mathsf{Conf})$, we get the following mu-calculus terms for $\mathit{Safe}[\emptyset](A)$ and $\mathit{Safe}[F](A)$:

- $\mathit{Safe}[\emptyset](A) = \nu X.\big(A \setminus \mathit{Pre}[\mathcal{F}](\mathsf{Conf})\big) \cap \widehat{Pre}_X(\mathsf{Conf})$, and
- $\mathit{Safe}[F](A) = \nu X.\big(A \setminus \mathit{Pre}[\mathcal{F} \setminus F](\mathsf{Conf})\big) \cap \bigcap_{f \in F} \widehat{Pre}_X^*(\widehat{Pre}_X[f](\mathsf{Conf}))$.

**Theorem 5.2 (Fair invariants).** *Let $A \subseteq \mathsf{Conf}$ and $s \in \mathsf{Conf}$.*

*(a) There is a $\mathcal{F}$-fair fm-scheduler $\mathcal{V}$ s.t. $\Pr_{\mathcal{V}}(s \models \Box A) > 0$ iff $s \models \exists(A \,\mathrm{Until}\, \mathit{Safe}_{\mathcal{F}}(A))$.*
*(b) There is a $\mathcal{F}$-fair fm-scheduler $\mathcal{V}$ s.t. $\Pr_{\mathcal{V}}(s \models \Box A) = 1$ iff $s \in \mathit{Prom}_A(\mathit{Safe}_{\mathcal{F}}(A))$.*

Observe that, for a region $\gamma$, $\mathit{Safe}_{\mathcal{F}}([\![\gamma]\!])$ and $\mathit{Prom}_A(\mathit{Safe}_{\mathcal{F}}([\![\gamma]\!]))$ are regions that can be built effectively (based on the same reasoning that we use for Theorem 3.2). Thus, Theorem 5.2 yields the decidability of the questions whether for a given NPLCS, region $A$ and configuration $s$, there exists a $\mathcal{F}$-fair fm-scheduler $\mathcal{U}$ such that $\Pr_{\mathcal{U}}(s \models \Box A) > 0$ or $= 1$.

In the sequel, for $A \subseteq \mathsf{Conf}$, we denote by $T^{\mathcal{F}}_{\Box A}$ the set of all configurations $s$ such that $\Pr_{\mathcal{U}}(s \models \Box A) = 1$ for some $\mathcal{F}$-fair fm-scheduler $\mathcal{U}$.

We now come to *repeated reachability* $\Box\Diamond A$ and *persistence* $\Diamond\Box A$ properties under fairness constraints. For $A \subseteq \mathsf{Conf}$, we define $T^{\mathcal{F}}_{\Box\Diamond A} = \bigcup_{F \subseteq \mathcal{F}} T_F$ where $T_F$ is the largest subset of $\mathsf{Conf} \setminus \mathsf{Enabl}(\mathcal{F} \setminus F)$ such that for all $t \in T_F$:

- there is a finite path $s_0 \xrightarrow{\delta_1} \cdots \xrightarrow{\delta_m} s_m$ with $m \geq 1$, $t = s_0$, $s_m \in A$ and $\mathit{Post}[\delta_i](s_{i-1}) \subseteq T_F$ for all $1 \leq i \leq m$,
- for each $f \in F$ there is a finite path $s_0 \xrightarrow{\delta_1} \cdots \xrightarrow{\delta_m} s_m$ with $t = s_0$, $m \geq 1$, $\delta_m \in f$ and $\mathit{Post}[\delta_i](s_{i-1}) \subseteq T_F$ for all $1 \leq i \leq m$.

**Theorem 5.3 (Fair repeated reachability and persistence).** *Let $A \subseteq \mathsf{Conf}$ and $s \in \mathsf{Conf}$.*

*(a) There exists a $\mathcal{F}$-fair fm-scheduler $\mathcal{U}$ with $\Pr_{\mathcal{U}}(s \models \Box\Diamond A) = 1$ iff $s \in \mathit{Prom}(T^{\mathcal{F}}_{\Box\Diamond A})$.*
*(b) There exists a $\mathcal{F}$-fair fm-scheduler $\mathcal{U}$ with $\Pr_{\mathcal{U}}(s \models \Box\Diamond A) > 0$ iff $s \in \mathit{Pre}^*(T^{\mathcal{F}}_{\Box\Diamond A})$.*
*(c) There exists a $\mathcal{F}$-fair fm-scheduler $\mathcal{U}$ with $\Pr_{\mathcal{U}}(s \models \Diamond\Box A) = 1$ iff $s \in \mathit{Prom}(T^{\mathcal{F}}_{\Box A})$.*

*(d) There exists a $\mathcal{F}$-fair fm-scheduler $\mathcal{U}$ with* $\Pr_{\mathcal{U}}(s \models \Diamond\Box A) > 0$ *iff* $s \in Pre^*(T^{\mathcal{F}}_{\Box A})$.

With similar arguments as for $Prom(A)$, the sets of configuration $T^{\mathcal{F}}_{\Box\Diamond A}$ and $T^{\mathcal{F}}_{\Box A} = Prom_A(Safe_{\mathcal{F}}(A))$ are regions whenever $A$ is a region. This entails the decidability of the questions whether given region $A$, there exists a $\mathcal{F}$-fair fm-scheduler $\mathcal{U}$ such that $\Pr_{\mathcal{U}}(s \models \varphi) = 1$ or $> 0$ where $\varphi = \Box\Diamond A$ or $\Diamond\Box A$.

We next consider *linear time* properties, formalized by LTL formulas $\varphi$ where regions serve as atomic propositions. The idea is to encode the fairness constraints in the model (the NPLCS) by a Streett property

$$fair = \bigwedge_{f\in\mathcal{F}} (\Box\Diamond A_f \to \Box\Diamond B_f)$$

(with regions $A_f, B_f \subseteq \mathsf{Conf}$) that will be considered in conjunction with $\varphi$. We modify the given LCS $\mathcal{L} = (Q, \mathsf{C}, \mathsf{M}, \Delta)$ and construct a new LCS $\mathcal{L}' = (Q', \mathsf{C}, \mathsf{M}, \Delta')$ as follows. We introduce new locations $q_F$ for all subsets $F$ of $\mathcal{F}$ and $q \in Q$, i.e., we deal with $Q' = \{q_F : q \in Q, F \subseteq \mathcal{F}\}$. $\Delta'$ is the smallest set of transition rules such that $p_G \xrightarrow{op} q_F \in \Delta'$ if $p \xrightarrow{op} q \in \Delta$, $G \subseteq \mathcal{F}$ and $F = \{f \in \mathcal{F} : p \xrightarrow{op} q \in f\}$. For $f \in \mathcal{F}$, $B_f$ is the set of configurations $\langle q_F, w\rangle$ in $\mathcal{L}'$ such that $f \in F$, while $A_f$ denotes the set of all configurations $\langle q_F, w\rangle$ of $\mathcal{L}'$ where $f$ is enabled in the configuration $\langle q, w\rangle$ of $\mathcal{L}$. We finally transform the given formula $\varphi$ into $\varphi'$ by replacing any region $C$ of $\mathcal{L}$ that appears as an atom in $\varphi$ with the region $C' = \{\langle q_F, w\rangle : \langle q, w\rangle \in C, F \subseteq \mathcal{F}\}$. For instance, if $\varphi = \Box\Diamond(q \wedge (c \neq \varepsilon))$ then $\varphi' = \Box\Diamond\big((q \vee \bigvee_{F\subseteq\mathcal{F}} q_F) \wedge (c \neq \varepsilon)\big)$.

In the sequel, let $\mathcal{N} = (\mathcal{L}, \tau)$ be the NPLCS that we want to verify against $\varphi$ and let $\mathcal{N}' = (\mathcal{L}', \tau)$ the associated modified NPLCS. Obviously, for each fm-scheduler $\mathcal{U}$ for $\mathcal{N}$ there is a "corresponding" fm-scheduler $\mathcal{U}'$ for $\mathcal{N}'$, and vice versa. Corresponding means that $\mathcal{U}'$ behaves as $\mathcal{U}$ for the current configuration $\langle q, w\rangle$ with $q \in Q$. If the current configuration of $\mathcal{U}'$ is $\langle q_F, w\rangle$ then $\mathcal{U}'$ behaves as $\mathcal{U}$ for $\langle q, w\rangle$. Then, $\Pr_{\mathcal{U}}(s \models \varphi) = \Pr_{\mathcal{U}'}(s \models \varphi')$ for all configurations $s$ in $\mathcal{N}$. Here, each configuration $s = \langle q, w\rangle$ of $\mathcal{N}$ is identified with the configuration $\langle q_\emptyset, w\rangle$ in $\mathcal{N}'$. Moreover, $\mathcal{U}$ is $\mathcal{F}$-fair iff $\Pr_{\mathcal{U}'}(s \models fair) = 1$. This yields part (a) of the following lemma. Part (b) follows from the fact that $\Pr_{\mathcal{U}}(s \models \varphi) = 1 - \Pr_{\mathcal{U}}(s \models \neg\varphi)$ for each scheduler $\mathcal{U}$.

**Lemma 5.4.** *Let $s$ be a configuration in $\mathcal{N}$ (and $\mathcal{N}'$) and $\varphi$ an LTL formula. Then:*

*(a) There exists a $\mathcal{F}$-fair fm-scheduler $\mathcal{U}$ for $\mathcal{N}$ such that* $\Pr_{\mathcal{U}}(s \models \varphi) = 1$ *if and only if there exists a fm-scheduler $\mathcal{U}'$ for $\mathcal{N}'$ such that* $\Pr_{\mathcal{U}'}(s \models fair \wedge \varphi') = 1$.

*(b) There exists a $\mathcal{F}$-fair fm-scheduler $\mathcal{U}$ for $\mathcal{N}$ such that* $\Pr_{\mathcal{U}}(s \models \varphi) = 0$ *if and only if there exists a fm-scheduler $\mathcal{V}$ for $\mathcal{N}'$ such that* $\Pr_{\mathcal{V}}(s \models fair \wedge \neg\varphi') = 1$.

*(c) There exists a $\mathcal{F}$-fair fm-scheduler $\mathcal{U}$ for $\mathcal{N}$ such that* $\Pr_{\mathcal{U}}(s \models \varphi) > 0$ *if and only if there exists a fm-scheduler $\mathcal{V}$ for $\mathcal{N}'$ such that* $\Pr_{\mathcal{V}}(s \models fair \wedge \varphi') > 0$.

*(d) There exists a $\mathcal{F}$-fair fm-scheduler $\mathcal{U}$ for $\mathcal{N}$ such that* $\Pr_{\mathcal{U}}(s \models \varphi) < 1$ *if and only if there exists a fm-scheduler $\mathcal{V}$ for $\mathcal{N}'$ such that* $\Pr_{\mathcal{V}}(s \models fair \wedge \neg\varphi') > 0$.

Lemma 5.4 even holds for arbitrary $\omega$-regular properties. It provides a reduction from the verification problem for qualitative LTL formulas in NPLCS's and fair fm-schedulers to the same problem for the full class of fm-schedulers. Thus, all decidability results that have been established for NPLCS's and qualitative verification problems for the class of fm-schedulers (see 4) also hold when fairness assumptions are made.

# 6 Automatic verification of Pachl's protocol

Fig. 1 directly translates into a LCS $\mathcal{L}_{\text{Pachl}}$ when the asynchronous product of $P_{\mathsf{L}}$ and $P_{\mathsf{R}}$ is considered. $\mathcal{L}_{\text{Pachl}}$ has $6\times 6 = 36$ control locations and $(18+18)\times 6 = 216$ transition rules. In order to reason about notions like "a rule $\delta$ has been fired", that are ubiquitous in fairness hypothesis, our tool adds an history variable recording the last fired rule (actually, only its action label). This would further multiply the number of states and of transitions by 20, but not all pairs (location,last action) are meaningful so that the final model can be stripped down to 144 locations and 948 rules. In all our results below we do not use the names of these 144 locations, but rather project them to the more readable underlying 36 locations.

## 6.1 Safety analysis

Pachl [22] computed manually the set $\mathit{Post}^*(\mathtt{Init})$ of all configurations reachable in $\mathcal{L}_{\text{Pachl}}$ from the initial empty configuration $\mathtt{Init} = (\mathtt{L0},\mathtt{R4},\varepsilon,\varepsilon)$, and such forward computations can sometimes be done automatically with the techniques described in [4] (although termination of the forward-reachability computations cannot be guaranteed in general). These computations show that the protocol does indeed preserve the integrity of communication in the sense that no confusion between data messages is introduced by losses.

Our calculus for regions is geared towards backward computation, where termination is guaranteed. Our implementation can compute automatically the set of deadlock configurations:

$$\mathtt{Dead} \stackrel{\text{def}}{=} \mathsf{Conf} \smallsetminus \mathit{Pre}(\mathsf{Conf}) \;=\; \langle \mathtt{L4},\mathtt{R4},\varepsilon,\varepsilon\rangle.$$

Hopefully, Dead is not reachable from Init. We can compute the set $\mathit{Pre}^*(\mathtt{Dead})$ of all unsafe configurations, that can end up in a deadlock. Intersecting with $\uparrow\mathtt{Init}$, we obtain the set of unsafe starting channel contents:

$$\mathit{Pre}^*(\mathtt{Dead}) \sqcap \uparrow\mathtt{Init} \;=\; \langle \mathtt{L0},\mathtt{R4},\uparrow\varepsilon,\uparrow\mathtt{a}_0\mathtt{d}_0\rangle \;+\; \langle \mathtt{L0},\mathtt{R4},\uparrow\mathtt{eod}\,\mathtt{a}_0,\uparrow\mathtt{a}_0\rangle \;+\; \langle \mathtt{L0},\mathtt{R4},\uparrow\mathtt{d}_0\mathtt{eod}\,\mathtt{a}_0,\uparrow\varepsilon\rangle.$$

Thus eventual deadlock is possible from location $(\mathtt{L0},\mathtt{R4})$ if the channels initially contain the appropriately unsafe contents.

## 6.2 Liveness analysis

We now come to what is the main motivation of our work: proving progress under fairness hypothesis. In this case study, the problem we address is in general to compute the set of all configurations satisfying some $\Pr_{\mathcal{U}}(s \models \Box\Diamond A) = 1$ for all schedulers $\mathcal{U}$ satisfying some fairness conditions $\mathcal{F}$. Following equivalences of section 5, this is related to the computation of $T^{\mathcal{F}}_{\Box\Diamond A}$. More precisely: $\{s \mid \forall \mathcal{U}\ \mathcal{F}\text{-fair } \Pr_{\mathcal{U}}(s \models \Box\Diamond A) = 1\} = \mathsf{Conf} \smallsetminus \mathit{Pre}^*(T^{\mathcal{F}}_{\Box\Diamond A})$.

When computing $T^{\mathcal{F}}_{\Box\Diamond A}$, all subsets of $\mathcal{F}$ have to be considered and this induces a combinatorial explosion for large $\mathcal{F}$. Since we did not yet develop and implement heuristics to overcome this difficulty, we only checked examples considering "small" $\mathcal{F}$ sets (meaning a number of fairness sets, each of which can be a large set of rules) in this preliminary study. For example, we considered "strong process fairness" $\mathcal{F}_{\texttt{process}} = \{F_{\texttt{left}}, F_{\texttt{right}}\}$ (with obvious meaning for the sets of transitions $F_{\texttt{left}}$, $F_{\texttt{right}}$), or "strong fairness for reading" $\mathcal{F}_{\texttt{read}} = \{F_{\texttt{read}}\}$.

Regarding the target set $A$, we consider questions whether a given transition (in $P_{\mathsf{L}}$ or $P_{\mathsf{R}}$) is fired infinitely often (using the history variable), or whether a process changes control states infinitely often, etc. Observe that a conjunction of "$\Pr_{\mathcal{U}}(s \models \Box\Diamond A_i) = 1$" gives $\Pr_{\mathcal{U}}(s \models \bigwedge_i \Box\Diamond A_i) = 1$, so that we can check formulas like $\bigwedge_i \Box\Diamond \texttt{Li} \wedge \bigwedge_i \Box\Diamond \texttt{Ri}$, expressing progress in communication between the two processes.

In the three following cases :

- $\mathcal{F} = \mathcal{F}_{\texttt{read}}$ and $A = \mathit{After}_{\texttt{left}}$
- $\mathcal{F} = \mathcal{F}_{\texttt{read}}$ and $A = \mathit{After}_{\texttt{left-move}}$
- $\mathcal{F} = \{F_{\texttt{read}}, F_{\texttt{right-read}}\}$ and $A = \mathit{After}_{\texttt{left}}$

our prototype model checker yields that $\texttt{Init} \in \mathsf{Conf} \setminus \mathit{Pre}^*(T^{\mathcal{F}}_{\Box\Diamond A})$. This means that, in all three cases, starting from Init, the set of configurations $A$ will be visited infinitely often almost surely, under all $\mathcal{F}$-fair schedulers.

## 7 Conclusion

We introduced NPLCS's, a model for nondeterministic channel systems where messages are lost probabilistically, and showed the decidability of qualitative verification question of the form "does $\varphi$ holds with probability 1 for all $\mathcal{F}$-fair finite-memory schedulers?" where $\varphi$ is an $\omega$-regular linear-time property and $\mathcal{F}$ a strong fairness condition.

When atomic propositions can refer to the contents of channels, which is required when one wants to express fairness and firability of rules, our decidability results rest upon a new notion of symbolic regions based on "prefixed upward-closures". These symbolic methods can be implemented rather directly and we used them to analyze simple systems.

These results are the outcome of a research project that started in [13, 14] with the first early definition of NPLCS's and was continued in [9] where the key notions for reducing to constrained reachability questions have been first identified in a simplified framework. Further developments will focus on incorporating algorithmic ideas from symbolic verification (normal forms, caches, sharing, ...) in our naive prototype verifier, turning it into a more solid analysis tool.

## References

1. P. A. Abdulla, C. Baier, S. Purushothaman Iyer, and B. Jonsson. Simulating perfect channels with probabilistic lossy channels. *Information and Computation*, 197(1–2):22–40, 2005.

2. P. A. Abdulla, N. Bertrand, A. Rabinovich, and Ph Schnoebelen. Verification of probabilistic systems with faulty communication. *Information and Computation*, 202(2):141–165, 2005.
3. P. A. Abdulla, A. Bouajjani, and J. d'Orso. Deciding monotonic games. In *Proc. 17th Int. Workshop Computer Science Logic (CSL 2003) and 8th Kurt Gödel Coll. (KGL 2003), Vienna, Austria, Aug. 2003*, volume 2803 of *Lecture Notes in Computer Science*, pages 1–14. Springer, 2003.
4. P. A. Abdulla, A. Collomb-Annichini, A. Bouajjani, and B. Jonsson. Using forward reachability analysis for verification of lossy channel systems. *Formal Methods in System Design*, 25(1):39–65, 2004.
5. P. A. Abdulla and B. Jonsson. Undecidable verification problems for programs with unreliable channels. *Information and Computation*, 130(1):71–90, 1996.
6. P. A. Abdulla and B. Jonsson. Verifying programs with unreliable channels. *Information and Computation*, 127(2):91–101, 1996.
7. C. Baier, N. Bertrand, and Ph. Schnoebelen. A note on the attractor-property of infinite-state Markov chains. *Information Processing Letters*, 97(2):58–63, 2006.
8. C. Baier, N. Bertrand, and Ph. Schnoebelen. On computing fixpoints in well-structured regular model checking, with applications to lossy channel systems. RR cs.CS/0606091, Computing Research Repository, June 2006. Visible at http://arxiv.org/abs/cs.CS/0606091.
9. C. Baier, N. Bertrand, and Ph. Schnoebelen. Verifying nondeterministic probabilistic channel systems against ω-regular linear-time properties. RR cs.LO/0511023, Computing Research Repository, April 2006. To be published in *ACM Trans. Computational Logic*, visible at http://arxiv.org/abs/cs.LO/0511023.
10. C. Baier and B. Engelen. Establishing qualitative properties for probabilistic lossy channel systems: An algorithmic approach. In *Proc. 5th Int. AMAST Workshop Formal Methods for Real-Time and Probabilistic Systems (ARTS '99), Bamberg, Germany, May 1999*, volume 1601 of *Lecture Notes in Computer Science*, pages 34–52. Springer, 1999.
11. C. Baier, B. R. Haverkort, H. Hermanns, J.-P. Katoen, and M. Siegle, editors. *Validation of Stochastic Systems – A Guide to Current Research*, volume 2925 of *Lecture Notes in Computer Science*. Springer, 2004.
12. C. Baier and M. Kwiatkowska. Model checking for a probabilistic branching time logic with fairness. *Distributed Computing*, 11(3):125–155, 1998.
13. N. Bertrand and Ph. Schnoebelen. Model checking lossy channels systems is probably decidable. In *Proc. 6th Int. Conf. Foundations of Software Science and Computation Structures (FOSSACS 2003), Warsaw, Poland, Apr. 2003*, volume 2620 of *Lecture Notes in Computer Science*, pages 120–135. Springer, 2003.
14. N. Bertrand and Ph. Schnoebelen. Verifying nondeterministic channel systems with probabilistic message losses. In Ramesh Bharadwaj, editor, *Proc. 3rd Int. Workshop on Automated Verification of Infinite-State Systems (AVIS 2004), Barcelona, Spain, Apr. 2004*, 2004.
15. D. Brand and P. Zafiropulo. On communicating finite-state machines. *Journal of the ACM*, 30(2):323–342, 1983.
16. C. Courcoubetis and M. Yannakakis. The complexity of probabilistic verification. *Journal of the ACM*, 42(4):857–907, 1995.
17. A. Finkel. Decidability of the termination problem for completely specificied protocols. *Distributed Computing*, 7(3):129–135, 1994.
18. B. Hailpern and S. Owicki. Verifying network protocols using temporal logic. In *Proc. NBS/IEEE Symposium on Trends and Applications 1980: Computer Network Protocols, Gaithersburg, MD, May 1980*, pages 18–28. IEEE Comp. Soc. Press, 1980.
19. S. Hart, M. Sharir, and A. Pnueli. Termination of probabilistic concurrent programs. *ACM Transactions on Programming Languages and Systems*, 5(3):356–380, 1983.
20. A. Kučera and Ph. Schnoebelen. A general approach to comparing infinite-state systems with their finite-state specifications. *Theoretical Computer Science*, 2006. To appear.

21. B. Masson and Ph. Schnoebelen. On verifying fair lossy channel systems. In *Proc. 27th Int. Symp. Math. Found. Comp. Sci. (MFCS 2002), Warsaw, Poland, Aug. 2002*, volume 2420 of *Lecture Notes in Computer Science*, pages 543–555. Springer, 2002.
22. J. K. Pachl. Protocol description and analysis based on a state transition model with channel expressions. In *Proc. 7th IFIP WG6.1 Int. Workshop on Protocol Specification, Testing, and Verification (PSTV '87), Zurich, Switzerland, May 1987*, pages 207–219. North-Holland, 1987.
23. P. Panangaden. Measure and probability for concurrency theorists. *Theoretical Computer Science*, 253(2):287–309, 2001.
24. A. Pnueli. On the extremely fair treatment of probabilistic algorithms. In *Proc. 15th ACM Symp. Theory of Computing (STOC '83), Boston, MA, Apr. 1983*, pages 278–290. ACM Press, 1983.
25. A. Pnueli and L. D. Zuck. Verification of multiprocess probabilistic protocols. *Distributed Computing*, 1(1):53–72, 1986.
26. A. Pnueli and L. D. Zuck. Probabilistic verification. *Information and Computation*, 103(1):1–29, 1993.
27. S. Purushothaman Iyer and M. Narasimha. Probabilistic lossy channel systems. In *Proc. 7th Int. Joint Conf. Theory and Practice of Software Development (TAPSOFT '97), Lille, France, Apr. 1997*, volume 1214 of *Lecture Notes in Computer Science*, pages 667–681. Springer, 1997.
28. M. L. Puterman. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 1994.
29. A. Rabinovich. Quantitative analysis of probabilistic lossy channel systems. In *Proc. 30th Int. Coll. Automata, Languages, and Programming (ICALP 2003), Eindhoven, NL, July 2003*, volume 2719 of *Lecture Notes in Computer Science*, pages 1008–1021. Springer, 2003.
30. Ph. Schnoebelen. Verifying lossy channel systems has nonprimitive recursive complexity. *Information Processing Letters*, 83(5):251–261, 2002.
31. Ph. Schnoebelen. The verification of probabilistic lossy channel systems. In Baier et al. [11], pages 445–465.
32. M. Y. Vardi. Automatic verification of probabilistic concurrent finite-state programs. In *Proc. 26th IEEE Symp. Foundations of Computer Science (FOCS '85), Portland, OR, USA, Oct. 1985*, pages 327–338. IEEE Comp. Soc. Press, 1985.