Reachability Analysis of Branching Probabilistic Processes

Emanuel Martinov

Doctor of Philosophy

Laboratory for Foundations of Computer Science
School of Informatics
University of Edinburgh

2021

Abstract

We study a fundamental class of infinite-state stochastic processes and stochastic games, namely Branching Processes, under the properties of (single-target) reachability and multi-objective reachability.

In particular, we study Branching Concurrent Stochastic Games (BCSGs), which are an imperfect-information game extension to the classical Branching Processes, and show that these games are determined, i.e., have a value, under the fundamental objective of reachability, building on and generalizing prior work on Branching Simple Stochastic Games and finite-state Concurrent Stochastic Games. We show that, unlike in the turn-based branching games, in the concurrent setting the almost-sure and limit-sure reachability problems do not coincide and we give polynomial time algorithms for deciding both almost-sure and limit-sure reachability. We also provide a discussion on the complexity of quantitative reachability questions for BCSGs.

Furthermore, we introduce a new model, namely Ordered Branching Processes (OBPs), which is a hybrid model between classical Branching Processes and Stochastic Context-Free Grammars. Under the reachability objective, this model is equivalent to the classical Branching Processes. We study qualitative multi-objective reachability questions for Ordered Branching Markov Decision Processes (OBMDPs), or equivalently context-free MDPs with simultaneous derivation. We provide algorithmic results for efficiently checking certain Boolean combinations of qualitative reachability and non-reachability queries with respect to different given target non-terminals.

Among the more interesting multi-objective reachability results, we provide two separate algorithms for almost-sure and limit-sure multi-target reachability for OBMDPs. Specifically, given an OBMDP, given a starting non-terminal, and given a set of target non-terminals, our first algorithm decides whether the supremum probability, of generating a tree that contains every target non-terminal in the set, is 1. Our second algorithm decides whether there is a strategy for the player to almost-surely (with probability 1) generate a tree that contains every target non-terminal in the set. The two separate algorithms are needed: we show that indeed, in this context, almost-sure and limit-sure multi-target reachability do not coincide. Both algorithms run in time polynomial in the size of the OBMDP and exponential in the number of targets. Hence, they run in polynomial time when the number of targets is fixed. The algorithms are fixed-parameter tractable with respect to this number. Moreover, we show that the qualitative almost-sure (and limit-sure) multi-target reachability decision problem is in general NP-hard, when the size of the set of target non-terminals is not fixed.

Lay Summary

The field of stochastic processes and stochastic games has been widely studied ever since the early to mid-twentieth century, with a wide range of applications across multiple disciplines. In the thesis, we investigate a well-known model, called Branching Processes, in this field and some specific extensions of it, and we ask questions that are fundamental and common to inquire about when a model in this field is investigated. Branching Processes are a classical class of stochastic processes, modelling the evolution of populations dependent on given probabilistic rules. Along with their specific extensions that we study, they are utilized as a modelling tool in areas, such as bioinformatics, biology, population genetics, physics and chemistry (e.g., chemical chain reactions), medicine (e.g., cancer growth), marketing and others.

In many cases, the process is not purely stochastic but there is the possibility of taking actions (e.g., adjusting the conditions of reactions, applying drug treatments in medicine, advertising in marketing, etc.) which can influence the probabilistic evolution of the process to bias it towards achieving desirable objectives. Some of the factors that affect the process may be controllable (to some extent) while others may not be sufficiently well-understood and thus it may be more appropriate to consider their affect in a probabilistic or in an adversarial manner. Some states in these processes are designated as (un)desirable (e.g., malignant cancer cells) and we may want to maximize or minimize the probability of reaching such states, where such a goal is generally referred to as the (single-target) reachability objective.

In the first half of the thesis, we study this (single-target) reachability objective for a specific extension of the model, where there are two players who simultaneously and independently of each other chose their actions and who have opposing goals, i.e., one aims to maximize the probability of reaching the specified state and the other to minimize it. In the second half of the thesis, we study another form of extensions to the model and an objective, which is certainly a natural extension to the (single-target) reachability objective.

Acknowledgements

First and foremost, I would like to thank my supervisor, Prof. Kousha Etessami, for his expertise and academic guidance throughout my doctorate programme, and for encouraging me to explore different areas of research in order to find my passion in academia. I have learned a lot from him both in terms of theoretical knowledge and about the essence of academic writing.

Many thanks to friends and fellow PhD students (they know who they are) for the countless hours and coffee sessions spent together inside and outside the office building chatting and keeping our spirits up.

Last but not the least, I would also like to thank my family for their moral support, for always being there to talk whenever I needed them and for continuously encouraging me to pursue my goals in life.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Emanuel Martinov*)

Publications

The following publications have appeared during the course of this doctorate:

Kousha Etessami, Emanuel Martinov, Alistair Stewart, Mihalis Yannakakis.
Reachability for Branching Concurrent Stochastic Games. In Proc. of 46th Int. Coll. on Automata, Languages and Programming (ICALP), 2019. [EMSY19]

Kousha Etessami, Emanuel Martinov.
Qualitative Multi-Objective Reachability for Ordered Branching MDPs. In Proc. of 14th Int. Conf. on Reachability Problems (RP), 2020. [EM20]

Table of Contents

List of Figures

List of Algorithms

Acronyms

MC	Markov Chain
MDP	Markov Decision Process
SSG	Simple Stochastic Game
CSG	Concurrent Stochastic Game
(O)BP	(Ordered) Branching Process
(O)BMDP	(Ordered) Branching MDP
(O)BSSG	(Ordered) Branching SSG
(O)BCSG	(Ordered) Branching CSG
SCFG	Stochastic Context-Free Grammar
RMC	Recursive Markov Chain
RMDP	Recursive MDP
RSSG	Recursive SSG
RCSG	Recursive CSG
(B)SCC	(Bottom) Strongly Connected Component
MEC	Maximal End-Component
CNF/DNF	Conjunctive Normal Form / Disjunctive Normal Form
SNF	Simple Normal Form
PPS	Probabilistic Polynomial System
LFP/GFP	Least Fixed Point / Greatest Fixed Point
LD(F)	Linear Degenerate (Free)
PH	Polynomial Hierarchy
CH	Counting Hierarchy
ETR ($\exists\mathbb{R}$)	Existential Theory of the Reals

Chapter 1

Introduction

The field of stochastic games is an extremely rich field, dating back to the 1950s with the introduction of finite-state zero-sum Concurrent (imperfect-information) Stochastic Games (CSGs) with discounted rewards by Shapley ([Sha53]). CSGs and their restricted subclasses of finite-state Simple (turn-based) Stochastic Games (SSGs) and Markov Decision Processes (MDPs) have been widely investigated through the years. The reason is the vast number of applications of these models in many disciplines, such as model verification, decision-theoretic planning, reinforcement learning, decision-making systems in the area of artificial intelligence and many others. The questions of the existence and type of (near-)optimal strategies, and the computational complexity of these models over various objectives have been well-studied, providing some nice techniques for solving them.

To name a few classical results, Shapley showed in [Sha53, Theorem 1] that computing the values (one for each start state) of a CSG can be represented as a fixed-point search problem over a specific system of equations (therefore, the same holds for the restricted subclasses). For the restricted subclass of SSGs, it has been shown that the decision problem of whether the reachability value is $\geq 1/2$ is in NP$\cap$coNP ([Con92, Theorem 1]), with both players having deterministic memoryless optimal strategies, and it is the well-known long-standing *Condon*'s open problem of whether it is decidable in P-time. Moreover, both the search problem of computing an exact value in Condon's SSGs and the search problem of computing an approximated value in Shapley's games are in PLS $\cap$ PPAD. Many problems in the field of stochastic games can be reduced to the *Condon*'s problem, or vice versa. Thus, its importance in the field is significant. In contrast, computing the reachability values in finite-state non-stochastic games is in P-time, based on graph-theoretic approach analysis ([Con92, Theorem 2]).

Computing the optimal reachability probabilities and the optimal deterministic memoryless strategy in (both maximizing and minimizing) finite-state 1-player MDPs can be done efficiently in P-time, by reducing the problem to solving Linear Programs. For more information, please refer to Puterman's book [Put94] on standard facts and theory for MDPs. Moreover, for a brief survey of well-known algorithms and techniques for solving MDPs and SSGs, please see, for instance, [Con93, Som05, LDK95].

As mentioned, one such technique is casting the solving of a MDP as a Linear Programming problem, where the latter is shown to be solvable in polynomial time in the size of the LP (and hence, in the size of the MDP) via the ellipsoid method approach by Khachiyan ([Kha79]) and latter via the more practical interior-point method approach by Karmarkar ([Kar84]). Another well-known technique is the *policy iteration or improvement*, often referred to as Hoffman-Karp algorithm ([HK66]), which involves improving players' strategies in an iterative manner and requires solving a LP at each iteration (applicable for solving SSGs as well). A third technique, also applicable for solving SSGs and widely-adapted to many other models including those studied in this thesis, is the *value iteration*, often referred to as *successive approximation*, introduced first in [Sha53], which is efficient within iterations but generally can take exponentially many rounds to achieve a constant factor approximation of the values (see [BKN$^+$19] for recent complexity analysis on value iteration). It involves the procedure of, starting in an initial feasible vector of values, repeatedly updating the values using a system of equations until the values vector converge to the optimal values vector in the limit.

In this thesis, we study fundamental objectives (properties) for certain infinite-state (but finitely represented) extensions of the aforementioned stochastic processes, namely we look at branching processes (and natural extensions of them) and discuss the properties of extinction/termination, (single-target) reachability and multi-objective reachability. In particular, we focus on the concurrent game generalization of Branching Processes and on the MDP (i.e., the 1-player) variant of Ordered Branching Processes, where the latter are stochastic processes that we have introduced in [EM20] (a paper that is incorporated in this thesis).

1.1 Branching Processes

Branching Processes (BPs) are a class of infinite-state stochastic processes that model the stochastic evolution of a population of objects of distinct types. In each generation, every object of each type, T, produces a *multi-set* of objects of various types in the next generation according to a given probability distribution on offsprings for the type T. BPs are a fundamental stochastic model that have been used to model phenomena in many fields, including bioinformatics and biology (see, e.g., [KA02]), population genetics ([HJV05]), physics and chemistry (e.g., particle systems, chemical chain reactions), medicine (e.g., cancer growth [Bea13, RBCN13]), marketing, and others. In many cases, the process is not purely stochastic but there is the possibility of taking actions (for example, adjusting the conditions of reactions, applying drug treatments in medicine, advertising in marketing, etc.) which can influence the probabilistic evolution of the process to bias it towards achieving desirable objectives. Some of the factors that affect the reproduction may be controllable (to some extent) while others are not and also may not be sufficiently well-understood to be modeled accurately by specific probability distributions, and thus it may be more appropriate to consider their effect in an adversarial (worst-case) sense. *Branching Concurrent Stochastic Games (BCSGs)* are a natural model to represent such settings. There are two players, who have a set of available actions for each type T that affect the reproduction for this type; for each object of type T in the evolution of the process, the two players select simultaneously and independently of each other an action from their available sets (possibly in a randomized manner) and their choice of actions determines the probability distribution for the offspring of the object. Therefore, BCSGs are imperfect-information zero-sum games. The first player represents the controller that can control some of the parameters of the reproduction and the second player represents other parameters that are not controlled and are treated adversarially. The first player wants to select a strategy that optimizes some objective. Some types are designated as undesirable (for example, malignant cells), in which case we want to minimize the probability of ever reaching any object of such a type. Or conversely, some types may be designated as desirable, in which case we want to maximize the probability of reaching an object of such a type. Hence, reachability is an essential objective to be studied in the model of branching processes.

BCSGs generalize the purely stochastic Branching Processes as well as Branching Markov Decision Processes (BMDPs) and Branching Simple Stochastic Games (BSSGs). In BMDPs there is only one player who aims to maximize or minimize the objective. In BSSGs there are two opposing players but they control different sets of types, i.e., the game is *turn-based* (perfect-information) zero-sum. These models were studied previously under the (single-target) *reachability* objective, namely the optimization of the probability of reaching a given target type ([ESY18]). They were also studied under another fundamental objective, namely the optimization of *extinction* probability, i.e., the probability that the process will eventually become extinct, that is, that the population will become empty ([ESY17, ESY20, EY09, EY15, EY08, EY06]). We will later (in Section 2.6) discuss in detail the prior results in these models and compare them with the results in this paper.

BCSGs can also be seen as a generalization of finite-state concurrent stochastic games (see [Eve57]), namely the extension of such finite games with branching. Concurrent games have been used in the verification area to model the dynamics of open systems, where one player represents the system and the other player the environment. Such a system moves sequentially from state to state depending on the actions of the two players (the system and the environment). Branching concurrent games model the more general setting in which processes can spawn new processes that proceed then independently in parallel (e.g., new threads are created and terminated).

The other model, which is a modification of the classical branching processes, that we introduce and study in this thesis is the model of *Ordered Branching Processes (OBPs)*.[1] Informally, one can think of OBPs as a hybrid model between Branching Processes and Stochastic Context-Free Grammars (SCFGs). And although it is formally defined in Section 2.4, in order to be slightly more precise here about how BPs and SCFGs are combined let us informally explain the 1-player-controlled generalization of OBPs, which is the main focus of Chapter 4. *Ordered Branching Markov Decision Processes (OBMDPs)* can be viewed as controlled/probabilistic context-free grammars, but without any terminal symbols, and where moreover the non-terminals are partitioned into two sets: controlled non-terminals and probabilistic non-terminals. Each non-terminal, N, has an associated set of grammar rules of the form $N \to \gamma$, where γ is a (possibly empty) sequence of non-terminals. Each probabilistic non-terminal is equipped with a given probability distribution on its associated grammar rules. For each controlled non-terminal, M, there is an associated non-empty set of available actions, A_M, which is in one-to-one correspondence with the grammar rules of M. So, for each action, $a \in A_M$, there is an associated grammar rule $M \xrightarrow{a} \gamma$. Given an OBMDP, given a "start" non-terminal, and given a "strategy" for the controller, these together determine a probabilistic process that generates a (possibly infinite) random ordered tree. The tree is formed via the usual parse tree expansion of grammar rules, proceeding generation by generation, in a top-down manner. Starting with a root node labeled by the "start" non-terminal, the ordered tree is generated based on the controller's

[1] OBPs were in fact introduced in [EM20], but that paper is a major part of this thesis.

(possibly randomized) choice of an action at each node of the tree that is labeled by a controlled non-terminal, and based on the probabilistic choice of a grammar rule at nodes that are labeled by a probabilistic non-terminal.

Ordered Branching Processes (OBPs) are OBMDPs without any controlled non-terminals. As mentioned, OBPs and their MDP and game generalizations are very similar to classical Branching Processes and their MDP and game generalizations, respectively. The difference is that for OBPs the generated tree is *ordered*. In particular, the rules for an OBP have an ordered *sequence* of non-terminals on their right hand side, whereas there is no such ordering in BPs: each rule for a given type associates an unordered multi-set of offsprings of various types to that given type.

In considering the functionality of OBPs, we have already covered the applications of BPs, but SCFGs also have well-known applications in many fields, including in natural language processing and RNA modeling ([DEKM98]). Generalizing these models to MDPs is natural, and can allow us to study, and to optimize algorithmically, settings where such random processes can partially be controlled.

It turns out that, under the (single-target) reachability and extinction/termination objectives, computing the (optimal) probabilities in BPs and OBPs (and similarly their MDP and game generalizations) is equivalent. However, this is not known for *multi-objective reachability*, which we focus on for OBMDPs in this thesis and leave as future work for BMDPs. Previously, multi-objective reachability has only been studied for the classical finite-state MDPs ([EKVY08]). We will also discuss in detail that under the (single-target) reachability and extinction/termination objectives, the models of BPs and OBPs (and similarly their MDP and game generalizations) are equivalent in some circumstances to other certain probabilistic processes. All these models are compared and contrasted in Section 2.6.

1.2 Major contributions and outline of the thesis

Chapter 2 provides all the necessary definitions and background on all the models and objectives discussed in the thesis. It includes previous related work and provides a survey of similarities and differences with other related stochastic processes with respect to the objectives that are the focus of the thesis. The chapter also shows that computing the (single-target) reachability probabilities and the extinction probabilities in BPs is equivalent to computing the (single-target) reachability probabilities and the termination probabilities, respectively, in OBPs. This way, Proposition 2.4 and subsection

2.4.1 act as a link between Chapters 3 and 4.

Chapter 3 contains the content of paper [EMSY19] and some further results not included in that paper. The chapter shows that computing the non-reachability values, starting at an object of any of the types, in BCSGs can be expressed as a system of equations (which we call a *minimax-PPS*), where there is a variable and an equation for each type and the right-hand side of each equation is the (von Neumann) minimax value of a zero-sum one-shot matrix game whose dimensions are defined by the available choices of both players in the particular type and whose entries are probabilistic polynomials. What is more, the chapter proves that the non-reachability values of the game are exactly the coordinates of the Greatest Fixed Point of the system. Next, the chapter shows that the qualitative almost-sure and limit-sure reachability problems do not coincide (unlike in the case of turn-based branching stochastic games) and provides polynomial time (in the size of the BCSG) algorithms for computing the types, starting at an object of which, almost-sure (respectively, limit-sure) reachability is achieved for the given target type. Here, the meaning of *achieved* is that the player maximizing the reachability probability has a strategy (respectively, a family of strategies) that guarantees almost-sure (respectively, limit-sure) reachability, regardless of the strategy of the player minimizing the reachability probability. The algorithms borrow techniques from [ESY18] and [dAHK07]. The proofs demonstrate how to compute an almost-sure strategy (respectively, a limit-sure strategy for a given error $\varepsilon > 0$) for the player maximizing the reachability probability, or alternatively, a spoiling strategy for the player minimizing the reachability probability if almost-sure (respectively, limit-sure) reachability is not satisfied.

Additionally, we adapt analogous results from [EY08, Theorem 3.3] and [EY09, Theorem 5.3] in order to make it clear to the reader that PSPACE is an upper bound for both quantitative reachability decision and approximation questions for BCSGs (this was previously known for the restricted subclass of BSSGs) and that PosSLP is a lower bound for the quantitative reachability decision questions even for the purely stochastic BPs (this was previously known for the extinction objective). These are the best bounds we know so far. We also show that computing exact optimal reachability probabilities for minimizing BMDPs is in the complexity class FIXP.

Chapter 4 contains the content of paper [EM20]. The chapter studies OBMDPs under a natural generalization of the standard reachability objective, namely multi-objective reachability where the player aims to optimize each of the respective probabilities that the generated tree satisfies each of several given objectives over different given target non-terminals. Our focus is on the qualitative multi-objective reachability decision questions, where a set of target non-terminals, K, is given and where, for each target non-terminal T_q ($q \in K$), we are also given a probability $b_q \in \{0,1\}$ and an inequality $\Delta_q \in \{<,=,>\}$, and where the goal is to decide, for any start non-terminal, whether the player has a single strategy using which, for all $q \in K$ the probability that the generated tree contains the non-terminal T_q is $\Delta_q b_q$. We provide efficient algorithms, i.e., running in time polynomial in the size of the OBMDP and in the number of targets, for deciding certain special cases of these problems.

But the most interesting results we provide are with respect to the qualitative multi-target reachability, i.e., the situation where a set of target non-terminals is given and, for a given starting non-terminal, the goal is to determine whether the player has a strategy to generate a tree that contains all targets almost-surely (or limit-surely). First, we give an example that demonstrates that, unlike for the standard single-target reachability objective, in the presence of even two targets almost-sure and limit-sure multi-target reachability do not coincide. We provide separate algorithms that compute the non-terminals, starting at which, almost-sure (respectively, limit-sure) multi-target reachability is achieved. We also provide an algorithm that computes the non-terminals, starting at which, regardless of the strategy there is a zero probability to generate a tree that contains all targets. We show that these problems are in general NP(coNP)-hard, when the number of given targets is unbounded. The provided algorithms for qualitative multi-target reachability questions run in time fixed-parameter tractable with respect to the number of targets and their proofs show how to construct the corresponding desired strategy for the player, e.g., a strategy that guarantees almost-sure multi-target reachability or a strategy that guarantees limit-sure multi-target reachability within a given desired error $\varepsilon > 0$.

Finally, Chapter 5 concludes by describing some of the open problems that we leave in this thesis, which provide interesting and promising future research.

Chapter 2

Background and Related Work

This chapter presents the necessary background and definitions for the models of Branching Processes and Ordered Branching Processes (and their MDP and game generalizations), and for the study of the problems analysed in Chapters 3 and 4. Furthermore, we show the similarities and differences between BPs and OBPs, and also to other closely-related stochastic processes, such as Stochastic Context-Free Grammars and Recursive Markov Chains. We also survey previous work on all these models with respect to the objectives studied in this thesis.

This chapter skips many standard definitions, which can be found in textbooks literature, such as Chung's book [Chu01] on probability theory and Puterman's book [Put94] on standard facts and theory for MDPs. Moreover, there is a vast amount of research and theory on Branching Processes, which is not covered in this thesis as it is not necessary. A good starting point is Harris's book [Har63].

Organization of the chapter. Section 2.1 recaps some important decision problems and complexity classes, that are referred to in the related work section and in the next chapters, in order to provide a better idea of where in the complexity hierarchy the analysed problems in this thesis reside. Sections 2.2 and 2.4 provide background required for the analysis of Branching Processes and Ordered Branching Processes, respectively. Section 2.3 introduces Probabilistic Polynomial Systems of equations, which are later (in Chapter 3) used to rephrase some analysed problems. Section 2.5 defines other related stochastic models, namely Stochastic Context-Free Grammars (2.5.1) and Recursive Markov models (2.5.2). Finally, Section 2.6 discusses past work related to (Ordered) Branching Processes and to the more general model of Recursive Markov chains.

2.1 Complexity

This section of the background chapter discusses some decision problems and complexity classes that will be referred to throughout this chapter and later chapters. We skip the definitions of other complexity classes which are also often referred to in the thesis as they are widely-known complexity classes. For general background on computational complexity, please refer to Arora and Barak's book [AB09], where especially relevant here are chapter 2 of the book on the complexity class NP and chapter 4 on space complexity (which includes the definition of PSPACE).

PosSLP and SQRT-SUM

Throughout the thesis we refer to the following two problems as important lower bounds for decision problems discussed in the thesis. PosSLP (*Positive Straight-Line Program*) is the problem of, given an arithmetic circuit C (equivalently, a straight-line program) with inputs 0 and 1 and over the basis of gates $\{+,-,*\}$, determining whether the output (i.e., the value from the top-most gate) is a positive number or not. It is a fundamental problem on arithmetic circuit complexity and it has been shown ([ABKPM09, Theorem 1.3]) to lie in the 4-th level of the *Counting Hierarchy* (CH) (i.e., PosSLP $\in P^{PP^{PP^{PP}}}$), which is the analog of the *Polynomial Hierarchy* (PH) for complexity classes for counting, such as #P. It is known that PH $\subseteq$ CH $\subseteq$ PSPACE.

The second problem, SQRT-SUM, is the problem of, given a collection of natural numbers $d_1,\ldots,d_n \in \mathbb{N}$ and another natural number $k \in \mathbb{N}$, determining whether $\sum_{i=1}^{n} \sqrt{d_i} \geq k$. It is a long-standing major open problem in the exact numerical computation complexity, not known to be in the Polynomial Hierarchy (not known to be even in NP, which was first set as a question in 1976 in a paper ([GGJ76]) about NP-complete geometric problems)[1]. It was shown in [ABKPM09, Proposition 1.1, Corollary 1.4] that SQRT-SUM is P-time reducible to PosSLP, hence placing it in the Counting Hierarchy.

Therefore, it is not believed that either of the two problems, PosSLP or SQRT-SUM, is PSPACE-hard, but placing them in PH would result in a major breakthrough on these long-standing problems.

[1] The version of the SQRT-SUM problem, where the comparison operator is $=$, is actually known to be in P-time ([Blo91]). There is a famous conjecture (see [Mal99, Proposition 1]) that the SQRT-SUM problem is efficiently decidable (i.e., in P-time), relying on the belief that it is enough to approximate each number $\sqrt{d_i}$ ($i \in [n]$) to polynomially many bits and then sum up the approximated numbers and compare to the given threshold number.

FIXP

There is a rich field of research on search problems that can be cast as fixed-point problems, i.e., problems where for each instance I of the search problem one can construct a continuous function F_I mapping a compact and convex domain D_I to itself, such that the set of solutions to the instance, $Sol(I)$, is exactly the set of fixed points of the function, $Fix(F_I)$. In this context, FIXP is the complexity class that captures search problems which can be rephrased as fixed-point problems for continuous functions, expressed by polynomial-size algebraic circuits (equivalently, straight-line programs) over the basis $\{+,-,*,/,\min,\max,\sqrt[k]{}\}$ with rational constants, over convex polytope domains described by linear inequalities and rational coefficients, where both the domain and the circuit can be computed in P-time in the size of the search-problem instance. This complexity class was introduced in [EY10], providing also the first FIXP-complete problem ([EY10, Theorem 18]), namely the computation of a Nash Equilibrium for 3 or more players ([EY10, Theorem 4] shows that this problem is both PosSLP-hard and SQRT-SUM-hard).

It was further shown in [EY10, Proposition 17] that, in contrast to the complexity class, FIXP, of real-valued search problems (where the complexity can be studied in a model of computation over the real numbers, such as the Blum-Shub-Smale (BSS) machine model [BSS89]), the corresponding discrete-valued complexity classes of decision problems (FIXP$_d$), approximation problems (FIXP$_a$) and Partial computation problems (FIXP$_{pc}$) are all contained in PSPACE, by relying on the upper bounds for decision procedures for the Existential Theory of the Reals (ETR or $\exists\mathbb{R}$) ([Ren92, Can88]). The ETR ($\exists\mathbb{R}$) decision problem is the problem of deciding whether a vector $x = (x_1,\ldots,x_n)$ exists that satisfies a given quantifier-free Boolean formula $\phi(x_1,\ldots,x_n)$, which consists of multi-variate polynomial inequalities and equalities with rational coefficients and over the variables $x = (x_1,\ldots,x_n)$. ETR is decidable in PSPACE ([Can88]) and in exponential time, where the exponent is a linear function of the number of variables ([Ren92]). To paint a better picture for the FIXP class, [EY10, Theorem 26] showed that if one is to restrict the basis for the algebraic circuits to the operations $\{+,-,\min,\max\}$, then this restricted complexity class (called, LINEAR-FIXP) is equal to the complexity class of total search problems, PPAD ([Pap94]), which lies between P and TFNP.

PPAD and PLS

Both complexity classes, PPAD and PLS, are subclasses of TFNP, i.e., of the class of total function problems solvable in non-deterministic polynomial time, and both classes have established an influential position in the complexity analysis of game theoretic problems. PPAD is a complexity class, introduced by Papadimitriou ([Pap94]), that captures some fixed-point problems, where a famous PPAD-complete problem is computing an exact Nash equilibrium for 2 players. PLS (*Polynomial Local Search*) is another search problem complexity class, that captures the complexity of finding a *local optimum* solution to an optimization problem. The crucial features of the PLS class are that for any problem residing in the class: there is a polynomial time computable function that returns the cost for each solution of an instance; and the neighbourhood of a solution in the domain can be searched in polynomial time, or in other words, one can verify that a solution is a local optimum or not in polynomial time.

These complexity classes also have an important place in the area of stochastic games. The search problem of computing the exact value of a Condon's SSG game lies in PPAD $\cap$ PLS (see [EY10, Corollary 25] and [Yan90]), and so is the search problem of computing the value of a Shapley's discounted concurrent stochastic game within a given desired error $\varepsilon > 0$, where the PPAD inclusion for the latter problem is proved in [EY10, Theorem 27] and the PLS inclusion for the latter problem follows from results in [EPRY20].

2.2 Branching Processes

This section introduces some definitions and background for Branching Concurrent Stochastic Games, generalizing some definitions in [ESY18] associated with reachability problems for Branching Markov Decision Processes and Branching Simple Stochastic Games.

We begin by defining the general model of Branching Concurrent Stochastic Games (BCSGs), as well as some important restrictions of the general model: Branching Simple Stochastic Games (BSSGs), Branching Markov Decision Processes (BMDPs), and Branching Processes (BPs).

Definition 1. *A* ***Branching Concurrent Stochastic Game (BCSG)*** *is a 2-player zero-sum game that consists of a finite set $V = \{T_1, \dots T_n\}$ of types, two finite non-empty sets $\Gamma^i_{max}, \Gamma^i_{min} \subseteq \Sigma$ of actions (one for each player) for each type T_i (Σ is a finite action alphabet), and a finite set $R(T_i, a_{max}, a_{min})$ of probabilistic rules associated with each tuple (T_i, a_{max}, a_{min}), $i \in [n]$, where $a_{max} \in \Gamma^i_{max}$ and $a_{min} \in \Gamma^i_{min}$. Each rule $r \in$*

$R(T_i, a_{max}, a_{min})$ is a triple (T_i, p_r, α_r), which we can denote by $T_i \xrightarrow{p_r} \alpha_r$, where $\alpha_r \in \mathbb{N}^n$ is a n-vector of natural numbers that denotes a finite multi-set over the set V, and where $p_r \in (0,1] \cap \mathbb{Q}$ is the probability of the rule r (which we assume to be a rational number, for computational purposes), where we assume that for all $T_i \in V$ and $a_{max} \in \Gamma^i_{max}$, $a_{min} \in \Gamma^i_{min}$, the rule probabilities in $R(T_i, a_{max}, a_{min})$ sum to 1, i.e., $\sum_{r \in R(T_i, a_{max}, a_{min})} p_r = 1$.

If for all types $T_i \in V$, either $|\Gamma^i_{max}| = 1$ or $|\Gamma^i_{min}| = 1$, then the model is a "turn-based" perfect-information game and is called a **Branching Simple Stochastic Game (BSSG)**. If for all $T_i \in V$, $|\Gamma^i_{max}| = 1$ (respectively, $|\Gamma^i_{min}| = 1$), then it is called a *minimizing* **Branching Markov Decision Process (BMDP)** (respectively, a *maximizing* BMDP). If both $|\Gamma^i_{min}| = 1 = |\Gamma^i_{max}|$ for all $i \in [n]$, then the process is a classic, purely stochastic, **multi-type Branching Process (BP)** ([Har63]).

A *play* of a BCSG defines a (possibly infinite) node-labeled forest, whose nodes are labeled by the type of the object they represent. A play contains a sequence of "generations", $X_0, X_1, X_2, \ldots$ (one for each integer time $t \geq 0$, corresponding to nodes at depth/level t in the forest). For each $t \in \mathbb{N}$, X_t consists of the population (a multi-set of objects of given types), at time t. X_0 is the initial population at generation 0 (these are the roots of the forest). X_{k+1} is obtained from X_k in the following way: for each object e in the population X_k, assuming e has type T_i, both players select simultaneously and independently actions $a_{max} \in \Gamma^i_{max}$, and $a_{min} \in \Gamma^i_{min}$ (or distributions on such actions), according to their strategies; thereafter a rule $r \in R(T_i, a_{max}, a_{min})$ is chosen randomly and independently (for object e) with probability p_r; each such object e in X_k is then replaced by the objects specified by the multi-set α_r associated with the corresponding randomly chosen rule r. This process is repeated in each generation, as long as the current generation is not empty, and if for some $k \geq 0$, $X_k = \emptyset$, then we say the process *terminates* or becomes *extinct*.

For a BCSG, the strategies of the players can in general be arbitrary. Specifically, at each generation, k, each player can, in principle, select actions for the objects in X_k based on the entire past history, may use randomization (a mixed strategy), and may make different choices for objects of the same type. The *history* of the process up to time $k-1$ is a forest of depth $k-1$ that includes not only the populations $X_0, X_1, \ldots, X_{k-1}$, but also the information regarding all the past actions and rules applied and the parent-child relationships between all the objects up to the generation of $k-1$. The history can be represented by a forest of depth $k-1$, with internal nodes labelled by rules and actions, and whose leaves at level $k-1$ form the population X_{k-1}.

Thus, a *strategy* of player 1 (player 2, respectively) is a function that maps every finite history (i.e., labelled forest of some finite depth as above) to a function that maps each object e in the current population X_k (assuming that the history has depth k) to a probability distribution on the actions Γ^i_{max} (on the actions Γ^i_{min}, respectively), assuming that object e has type T_i.

Let Ψ_1, Ψ_2 be the set of all strategies of players 1, 2, respectively. We say that a strategy is *deterministic* if for every history it maps each object e in the current population to a single action with probability 1 (in other words, it does not randomize on actions). We say that a strategy is *static* if for each type $T_i \in V$, and for any object e of type T_i, the player always chooses the same distribution on actions, irrespective of the history. That is, a *static* strategy is not only *memoryless* (i.e., does not depend on past history), but also uses the same distribution on actions for any two objects of the same type that reside in the same generation.

Different objectives can be considered for the BCSG game model. To name two, that are fundamental and are discussed in this thesis:

- *extinction* objective, where the aim of the players is to maximize/minimize the extinction probability, i.e., the probability of reaching a generation $X_k = \emptyset$, $k \geq 0$.
- *(single-target) reachability* objective (the focus of Chapter 3), where the aim of the players is to maximize/minimize the probability of reaching a generation X_k, $k \geq 0$, that contains at least one object of a given target type T_{f^*}.

Let us note right away that there is a natural "duality" between the objectives of optimizing reachability probability and that of optimizing extinction probability for branching processes. This duality was previously detailed in [ESY18] for BSSGs. The objective of optimizing the extinction probability of a BCSG, starting with an object of a given type, can equivalently be rephrased as a "*universal reachability*" objective (on a slightly modified BCSG), where the goal is to optimize the probability of eventually reaching the target type on *all* paths starting at the root of the tree. To see this, consider a modified BCSG with a target type, called *death*, and where for every type T_i, every rule $T_i \xrightarrow{p_r} \emptyset$ in the original BCSG is replaced with rule $T_i \xrightarrow{p_r} \mathit{death}$ in the modified BCSG. Likewise, the "universal reachability" objective can be rephrased as the objective of optimizing the extinction probability (on a slightly modified BCSG). To be more specific, consider a modified BCSG where for every type T_i, every rule $T_i \xrightarrow{p_r} \alpha_r$, $\alpha_r \neq \emptyset$, in the original BCSG is replaced by the rule $T_i \xrightarrow{p_r} \alpha'_r$ in the modified BCSG such that α'_r is the same as α_r but instead all copies of the target type are removed. Also, a new probabilistic type *dead* is introduced with a single rule $dead \xrightarrow{1} dead$, and for every non-target type T_i and every rule $T_i \xrightarrow{p_r} \emptyset$ in the original BCSG is replaced by $T_i \xrightarrow{p_r} dead$ in the modified BCSG.

By contrast, the *reachability* objective that we study in Chapter 3 is the "*existential reachability*" objective of optimizing the probability of reaching the target type on *some* path in the generated tree.

Despite this natural duality between the objectives of reachability and extinction, there is a wide disparity between them, both in terms of the nature and existence of optimal strategies, and in terms of computational complexity. For detailed past related work on these objectives with respect to BPs and related models, see Section 2.6.

The BCSG reachability game can of course also be viewed as a "non-reachability" game (by just reversing the roles of the players). It turns out this is useful to do, and we will exploit it in crucial ways (and this was also exploited in [ESY18] for BMDPs and BSSGs). So we provide some notation for this purpose.

Given an initial population $\mu \in \mathbb{N}^n$, with $\mu_{f^*} = 0$, and given an integer $k \geq 0$, and strategies $\sigma \in \Psi_1, \tau \in \Psi_2$, let $g^k_{\sigma,\tau}(\mu)$ be the probability that the process does *not* reach a generation with an object of type T_{f^*} in at most k steps, under strategies σ, τ and starting from the initial population μ. To be more formal, this is the probability that $(X_l)_{f^*} = 0$ for all $0 \leq l \leq k$. Similarly, let $g^*_{\sigma,\tau}(\mu)$ be the probability that $(X_l)_{f^*} = 0$ for all $l \geq 0$. We define $g^k(\mu) = \sup_{\sigma \in \Psi_1} \inf_{\tau \in \Psi_2} g^k_{\sigma,\tau}(\mu)$ to be the *value* of the k-step non-reachability game for the initial population μ, and $g^*(\mu) = \sup_{\sigma \in \Psi_1} \inf_{\tau \in \Psi_2} g^*_{\sigma,\tau}(\mu)$ to be the *value* of the game under the non-reachability objective and for the initial population μ. Section 3.1 demonstrates that these games are determined, meaning they have a value where $g^*(\mu) = \sup_{\sigma \in \Psi_1} \inf_{\tau \in \Psi_2} g^*_{\sigma,\tau}(\mu) = \inf_{\tau \in \Psi_2} \sup_{\sigma \in \Psi_1} g^*_{\sigma,\tau}(\mu)$. This implies that for every $\varepsilon > 0$, the player maximizing (minimizing) the non-reachability probability has a strategy to guarantee probability $\geq g^*(\mu) - \varepsilon$ (respectively, $\leq g^*(\mu) + \varepsilon$), regardless of what the other player does. Similarly, for $g^k(\mu)$.

In the case where the initial population μ is a single object of some given start type T_i[2], then for the value of the game we write g^*_i (or similarly, g^k_i), and when strategies σ and τ are fixed, we write $(g^*_{\sigma,\tau})_i$. The collection of these values, namely the vector g^* of g^*_i's, is called the *vector of the non-reachability values* of the game. We will see that, having the vector of g^*_i's, the non-reachability value for a starting population

[2]We can assume w.l.o.g. that the initial population consists of a single object of some given type T_i, because for any initial population $\mu \in \mathbb{N}^n$ of multiple objects, we can always add an auxiliary type T_j to the set V, where $\Gamma^j_{max} = \{a\} = \Gamma^j_{min}$ and the set $R(T_j, a, a)$ consists of a single probabilistic rule $T_j \xrightarrow{1} \mu$.

μ can be computed simply as $g^*(\mu) = f(g^*,\mu) := \prod_i (g_i^*)^{(\mu)_i}$ (see Section 3.1). So given a BCSG, the *aim* is to compute the vector of non-reachability values. As our original objective is reachability, we point out that the vector of reachability values is $r^* = \mathbf{1} - g^*$ (where $\mathbf{1}$ is the all-1 vector), and hence the reachability game value $r^*(\mu)$, starting with population μ, is $r^*(\mu) = 1 - g^*(\mu)$.

We study both qualitative and quantitative problems for the (non-)reachability objective in BCSGs. Let us define the problems in terms of the provided notation for non-reachability probabilities and values. The *qualitative almost-sure reachability* problem is the question of deciding, starting with an object of some given type T_i, whether there exists a strategy $\tau^* \in \Psi_2$ for the player minimizing the non-reachability probability (i.e., maximizing the reachability probability) such that $(g^*_{*,\tau^*})_i = 0$. The *qualitative limit-sure reachability* problem is the question of deciding, starting with an object of some given type T_i, whether $g_i^* = 0$, or in other words, whether for every $\varepsilon > 0$ there is a strategy $\tau_\varepsilon \in \Psi_2$ such that $\forall \sigma \in \Psi_1 : \ (g^*_{\sigma,\tau_\varepsilon})_i \leq \varepsilon$. The quantitative problems divide into decision and approximation problems. The *quantitative reachability decision* problem is the question of deciding, starting with an object of some given type T_i and given some rational value $p \in [0,1]$, whether $g_i^* \triangle p$, where $\triangle \in \{<,\leq,=,>,\geq\}$. The *quantitative reachability approximation* problem is the problem of, starting with an object of some given type T_i and given a desired error $\varepsilon > 0$, computing a value v such that $|g_i^* - v| \leq \varepsilon$.

Finally, note that any Branching Process, $\mathcal{A}$, defines a global infinite-state Markov chain, $M^{\mathcal{A}} = (Q,\Delta)$, where the global states Q are labeled finite trees, $\mathcal{T}$ (i.e., each global state is a finite sequence of generations $X_0, X_1, \ldots, X_t$, $t \geq 0$), and a transition $(\mathcal{T}, p_{\mathcal{T},\mathcal{T}'}, \mathcal{T}') \in \Delta$ exists for global states $\mathcal{T}, \mathcal{T}' \in Q$ if and only if there is a sequence of rules, $\beta = \langle r_1, \ldots, r_z \rangle$, such that tree $\mathcal{T}'$ can be obtained from tree $\mathcal{T}$ in one generation step using β (i.e., such that $\mathcal{T}$ is a prefix of $\mathcal{T}'$ and, if the last generation in tree $\mathcal{T}$ consists of objects of types $\langle T_{i_1}, T_{i_2}, \ldots, T_{i_z} \rangle$, then the last generation in tree $\mathcal{T}'$ consists of the objects of the types given by the collection of multi-sets $\alpha_{r_1}, \alpha_{r_2}, \ldots, \alpha_{r_z}$, where for every $j \in [z]$: there exists a rule $r_j \in R(T_{i_j})$ that satisfies $T_{i_j} \xrightarrow{p_{r_j}} \alpha_{r_j}$). The probability of the transition is $p_{\mathcal{T},\mathcal{T}'} := \prod_{j \in [z]} p_{r_j}$.

2.3 Systems of Probabilistic Polynomial Equations

We will later (in Section 3.1) show how to associate with any given BCSG a system of *minimax probabilistic polynomial equations (minimax-PPS)*, $x = P(x)$, for the non-reachability objective. This system will be constructed to have one variable x_i and one equation $x_i = P_i(x)$ for each type T_i other than the target type T_{f^*}.

In order to define these systems of equations, some shorthand notation will be useful. We use x^v to denote the monomial $x_1^{v_1} * x_2^{v_2} \cdots * x_n^{v_n}$ for an n-vector of variables $x = (x_1, \cdots, x_n)$ and a vector $v \in \mathbb{N}^n$. Considering a multi-variate polynomial $P_i(x) = \sum_{r \in R} p_r x^{\alpha_r}$ for some rational coefficients $p_r, r \in R$, we will call $P_i(x)$ a **probabilistic polynomial**, if $p_r \geq 0$ for all $r \in R$ and $\sum_{r \in R} p_r \leq 1$.

Definition 2. *A **probabilistic polynomial system of equations (PPS)**, $x = P(x)$, is a system of n equations, $x_i = P_i(x)$, in n variables where for all $i \in \{1, \ldots, n\}$, $P_i(x)$ is a probabilistic polynomial.*

*A **minimax probabilistic polynomial system of equations (minimax-PPS)**, $x = P(x)$, is a system of n equations in n variables $x = (x_1, \ldots, x_n)$, where for each $i \in \{1, \ldots, n\}$, there is an associated* MINIMAX-PROBABILISTIC-POLYNOMIAL $P_i(x) := Val(A_i(x))$. *By this we mean that $P_i(x)$ is defined to be, for each $x \in \mathbb{R}^n$, the minimax value of the two-player zero-sum matrix game given by a finite game payoff matrix $A_i(x)$ whose rows are indexed by the actions Γ^i_{max}, and whose columns are indexed by the actions Γ^i_{min}, where, for each pair $a_{max} \in \Gamma^i_{max}$ and $a_{min} \in \Gamma^i_{min}$, the matrix entry $A_i(x)_{(a_{max}, a_{min})}$ is given by a probabilistic polynomial $q_{i, a_{max}, a_{min}}(x)$. Thus, if $n_i = |\Gamma^i_{max}|$ and $m_i = |\Gamma^i_{min}|$, and if we assume w.l.o.g. that $\Gamma^i_{max} = \{1, \ldots, n_i\}$ and that $\Gamma^i_{min} = \{1, \ldots, m_i\}$, then $Val(A_i(x))$ is defined as the minimax value of the zero-sum matrix game, given by the following payoff matrix:*

$$A_i(x) = \begin{bmatrix} q_{i,1,1}(x) & q_{i,1,2}(x) & \ldots & q_{i,1,m_i}(x) \\ q_{i,2,1}(x) & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ \vdots & \vdots & \vdots & \vdots \\ q_{i,n_i,1}(x) & \ldots\ldots & \ldots\ldots & q_{i,n_i,m_i}(x) \end{bmatrix}$$

with each $q_{i,j,k}(x) := \sum_{r \in R(T_i,j,k)} p_r x^{\alpha_r}$ being a probabilistic polynomial for the actions pair j,k.

*If for all $i \in \{1, \ldots, n\}$, either $|\Gamma^i_{min}| = 1$ or $|\Gamma^i_{max}| = 1$, then we call such a system **min-max-PPS**. If for all $i \in \{1, \ldots, n\}$, $|\Gamma^i_{min}| = 1$ (respectively, if $|\Gamma^i_{max}| = 1$ for all i) then we will call such a system a **maxPPS** (respectively, a **minPPS**). Finally, a **PPS** is a minimax-PPS with both $|\Gamma^i_{min}| = 1 = |\Gamma^i_{max}|$ for every $i \in \{1, \cdots, n\}$.*

For computational purposes, we assume that all coefficients are rational and that there are no zero terms in the probabilistic polynomials, and we assume the coefficients and non-zero exponents of each term are given in binary. We denote by $|P|$ the total bit encoding length of a system, $x = P(x)$, under this representation.

As it will be later discussed, since $P(\cdot)$ defines a monotone function $P : [0,1]^n \rightarrow [0,1]^n$, it follows by Tarski's theorem ([Tar55, Theorem 1]) that any such system has both a **Least Fixed Point (LFP)** solution, $q^* \in [0,1]^n$, and a **Greatest Fixed Point(GFP)** solution, $g^* \in [0,1]^n$. In other words, $q^* = P(q^*)$ and $g^* = P(g^*)$ and moreover, for any $s^* \in [0,1]^n$ such that $s^* = P(s^*)$, we have $q^* \leq s^* \leq g^*$ (coordinate-wise inequality).

Definition 3. *A (possibly randomized)* ***policy*** *for the max (min) player in a minimax-PPS, $x = P(x)$, is a function that assigns a probability distribution to each variable x_i such that the support of the distribution is a subset of Γ^i_{max} (Γ^i_{min}, respectively), where these now denote the possible actions (i.e., choices of rows and columns) available for the respective player in the matrix game $A_i(x)$ that defines $P_i(x)$.*

Intuitively, a policy is the same as a static strategy in the corresponding BCSG.

Definition 4. *For a minimax-PPS, $x = P(x)$, and policies σ and τ for the max and min players, respectively, we write $x = P_{\sigma,\tau}(x)$ for the PPS obtained by fixing both these policies. We write $x = P_{\sigma,*}(x)$ for the minPPS obtained by fixing σ for the max player, and $x = P_{*,\tau}(x)$ for the maxPPS obtained by fixing τ for the min player. More specifically, for policy σ for the max player, we define the minPPS, $x = P_{\sigma,*}(x)$, as follows: for all $i \in [n]$, $(P_{\sigma,*}(x))_i := \min\{s_k : k \in \Gamma^i_{min}\}$, where $s_k := \sum_{j \in \Gamma^i_{max}} \sigma(x_i, j) * q_{i,j,k}(x)$, where $\sigma(x_i, j)$ is the probability that the fixed policy σ assigns to action $j \in \Gamma^i_{max}$ in variable x_i. We similarly define $x = P_{*,\tau}(x)$ and $x = P_{\sigma,\tau}(x)$.*

For a minimax-PPS, $x = P(x)$, and a (possibly randomized) policy σ for the max player, we use $q^*_{\sigma,*}$ and $g^*_{\sigma,*}$ to denote the LFP and GFP solution vectors of the corresponding minPPS, $x = P_{\sigma,*}(x)$, respectively. Likewise we use $q^*_{*,\tau}$ and $g^*_{*,\tau}$ to denote the LFP and GFP solution vectors of the maxPPS, $x = P_{*,\tau}(x)$, and we use $q^*_{\sigma,\tau}$ and $g^*_{\sigma,\tau}$ to denote the LFP and GFP solution vectors of the PPS, $x = P_{\sigma,\tau}(x)$.

Note: we overload notations such as $(g^*_{\sigma,*})_i$ and $(g^*_{*,\tau})_i$ to mean slightly different things, depending on whether σ and τ are static strategies (policies), or are more general non-static strategies. Specifically, let $E_i \in \mathbb{N}^n$ denote the unit vector which is 1 in the i-th coordinate and 0 elsewhere. When $\tau \in \Psi_2$ is a general non-static strategy we use the notation $(g^*_{*,\tau})_i := g^*_{*,\tau}(E_i) = \sup_{\sigma \in \Psi_1} g^*_{\sigma,\tau}(E_i)$, i.e., $(g^*_{*,\tau})_i$ will denote the optimal non-reachability probability starting with one object of type T_i and under fixed strategy τ for the min player. We likewise define $(g^*_{\sigma,*})_i$.

If, however, τ is a static strategy (policy), then $(g^*_{*,\tau})_i$ will denote both the aforementioned and the i-th coordinate in the GFP of maxPPS, $x = P_{*,\tau}(x)$, which as later discussed happen to be the same thing. Similarly, if σ is static. It will typically be clear from the context which interpretation of $(g^*_{*,\tau})_i$ is intended.

Definition 5. *For a minimax-PPS, $x = P(x)$, a policy σ^* is called* ***optimal*** *for the max player for the LFP (respectively, the GFP) if $q^*_{\sigma^*,*} = q^*$ (respectively, $g^*_{\sigma^*,*} = g^*$).*

An optimal policy τ^* for the min player for the LFP and GFP, respectively, is defined similarly.

For $\varepsilon > 0$, a policy σ' for the max player is called ***ε-optimal*** *for the LFP (respectively, the GFP), if $||q^*_{\sigma',*} - q^*||_\infty \leq \varepsilon$ (respectively, $||g^*_{\sigma',*} - g^*||_\infty \leq \varepsilon$). An ε-optimal policy τ' for the min player is defined similarly.*

For convenience in proofs throughout the thesis and to simplify the structure of the matrices involved in the *minimax-probabilistic-polynomials*, $P_i(x)$, we shall observe that minimax-PPSs can always be cast in the following normal form.

Definition 6. *A minimax-PPS in* ***simple normal form (SNF)****, $x = P(x)$, is a system of n equations in n variables $\{x_1, \cdots, x_n\}$, where each $P_i(x)$ for $i = 1, 2, \ldots, n$ is one of three forms:*

- FORM L: $P_i(x) = a_{i,0} + \sum_{j=1}^n a_{i,j} x_j$, *where for all j, $a_{i,j} \geq 0$, and $\sum_{j=0}^n a_{i,j} \leq 1$.*
- FORM Q: $P_i(x) = x_j x_k$ *for some j, k.*
- FORM M: $P_i(x) = Val(A_i(x))$, *where $A_i(x)$ is a $(n_i \times m_i)$ matrix, such that for all $a_{max} \in [n_i]$ and $a_{min} \in [m_i]$, the entry $A_i(x)_{(a_{max}, a_{min})} \in \{x_1, \ldots, x_n\} \cup \{1\}$.*

  (The reason we also allow "1" as an entry in the matrices $A_i(x)$ will become clear later in the context of the algorithms.)

We shall often assume a minimax-PPS in its SNF form, and say that a variable x_i is "of form/type" L, Q, or M, meaning that $P_i(x)$ has the corresponding form. The following proposition shows that we can efficiently convert any minimax-PPS into its SNF-form.

Proposition 2.1 (cf. [EY09, ESY18])**.** *Every minimax-PPS, $x = P(x)$, can be transformed in P-time to an "equivalent" minimax-PPS, $y = Q(y)$, in SNF form, such that $|Q| \in O(|P|)$. More precisely, the variables x are a subset of the variables y, and both* *the LFP and GFP of $x = P(x)$ are, respectively, the projection of the LFP and GFP of $y = Q(y)$, onto the variables x, and furthermore an optimal (respectively, ε-optimal) policy for the LFP (respectively, GFP) of $x = P(x)$ can be obtained in P-time from an optimal (respectively, ε-optimal) policy for the LFP (respectively, GFP) of $y = Q(y)$.*

Proof. We can easily convert, in P-time, any minimax-PPS into SNF form, using the following procedure.

- For each equation $x_i = P_i(x) := Val(A_i(x))$, for each probabilistic polynomial $q_{i,j,k}(x)$ on the right-hand side that is not a variable, add a new variable x_d, replace $q_{i,j,k}(x)$ with x_d in $P_i(x)$, and add the new equation $x_d = q_{i,j,k}(x)$.

- For each equation $x_i = P_i(x) = \sum_{j=1}^{m} p_j x^{\alpha_j}$, where $P_i(x)$ is a probabilistic polynomial that is not just a constant or a single monomial, replace every (non-constant) monomial x^{α_j} on the right-hand side that is not a single variable by a new variable x_{i_j} and add the equation $x_{i_j} = x^{\alpha_j}$.

- For each variable x_i that occurs in some polynomial with exponent higher than 1, introduce new variables $x_{i_1}, \ldots, x_{i_k}$ where k is the logarithm of the highest exponent of x_i that occurs in $P(x)$, and add equations $x_{i_1} = x_i^2$, $x_{i_2} = x_{i_1}^2, \ldots, x_{i_k} = x_{i_{k-1}}^2$. For every occurrence of a higher power x_i^l, $l > 1$, of x_i in $P(x)$, if the binary representation of the exponent l is $a_k \ldots a_2 a_1 a_0$, then we replace x_i^l by the product of the variables x_{i_j} such that the corresponding bit a_j is 1, and x_i if $a_0 = 1$. After we perform this replacement for all the higher powers of all the variables, every polynomial of total degree > 2 is just a product of different variables.

- If a polynomial $P_i(x) = x_{j_1} \ldots x_{j_m}$ in the current system is the product of $m > 2$ variables, then add $m-2$ new variables $x_{i_1}, \ldots, x_{i_{m-2}}$, set $P_i(x) = x_{j_1} x_{i_1}$, and add the equations $x_{i_1} = x_{j_2} x_{i_2}$, $x_{i_2} = x_{j_3} x_{i_3}, \ldots, x_{i_{m-2}} = x_{j_{m-1}} x_{j_m}$.

Now all equations are of the form L, Q, or M.

The above procedure allows us to convert any minimax-PPS, $x = P(x)$, into one in SNF-form by introducing $O(|P|)$ new variables and blowing up the size of P by a constant factor $O(1)$. It is clear that both the LFP and the GFP of $x = P(x)$ arise as the projections of the LFP and GFP of $y = Q(y)$ onto the x variables. Furthermore, there is an obvious (and easy to compute) bijection between policies for the resulting SNF-form minimax-PPS and the original minimax-PPS. □

Thus from now on, and for the rest of this thesis unless explicitly specified, *we may assume if needed, without loss of generality, that all minimax-PPSs are in SNF normal form.*

Definition 7. *The **dependency graph** of a minimax-PPS, $x = P(x)$, is a directed graph that has one node x_i for each variable x_i, and contains an edge (x_i, x_j) if x_j appears in $P_i(x)$. The dependency graph of a BCSG has one node T_i for each type T_i, and contains an edge (T_i, T_j) if there is a pair of actions $a_{max} \in \Gamma^i_{max}, a_{min} \in \Gamma^i_{min}$ and a rule $T_i \xrightarrow{p_r} \alpha_r$ in $R(T_i, a_{max}, a_{min})$ such that $(\alpha_r)_j \geq 1$.*

2.4 Ordered Branching Processes

This section introduces background and definitions for Ordered Branching Markov Decision Processes (OBMDPs) and the restricted model of Ordered Branching Processes (OBPs), and for the analysis of multi-objective reachability in these models. First, we define OBMDPs in a general way that combines both control and probabilistic rules at each non-terminal, and that allows rules to have an arbitrarily-long string of non-terminals on their right-hand side. Then we show that any OBMDP can be converted efficiently to an "equivalent"[3] one in "normal" form.

Definition 8. *An **Ordered Branching Markov Decision Process (OBMDP)**, $\mathcal{A}$, is a 1-player controlled stochastic process, represented by a tuple $\mathcal{A} = (V, \Sigma, \Gamma, R)$, where $V = \{T_1, \dots, T_n\}$ is a finite set of non-terminals, and Σ is a finite non-empty action alphabet. For each $i \in [n]$, $\Gamma^i \subseteq \Sigma$ is a finite non-empty set of actions for non-terminal $T_i \in V$, and for each $a \in \Gamma^i$, $R(T_i, a)$ is a finite set of probabilistic rules associated with the pair (T_i, a). Each rule $r \in R(T_i, a)$ is a triple, denoted by $T_i \xrightarrow{p_r} s_r$, where $s_r \in V^*$ is a (possibly empty) ordered sequence (string) of non-terminals and $p_r \in (0, 1] \cap \mathbb{Q}$ is the positive probability of the rule r (which we assume to be a rational number for computational purposes). We assume that for each non-terminal $T_i \in V$ and each $a \in \Gamma^i$, the rule probabilities in $R(T_i, a)$ sum to 1, i.e., $\sum_{r \in R(T_i, a)} p_r = 1$.*

We denote by $|\mathcal{A}|$ the total bit encoding length of the OBMDP. If $|\Gamma^i| = 1$ for all non-terminals $T_i \in V$, then the model is called an **Ordered Branching Process (OBP)**. Adding a second player (as an adversary), similarly to Section 2.2, we obtain an **Ordered Branching Simple (i.e., turn-based) Stochastic Game (OBSSG)**

[3] Equivalent w.r.t. all (multi-objective) reachability objectives we consider.

or an **Ordered Branching Concurrent Stochastic Game (OBCSG)**, depending on whether, respectively, the two players control disjoint sets of non-terminals or they both simultaneously and independently control each non-terminal.

In order to simplify the structure of the OBMDP model and to facilitate the proofs throughout the paper, we observe a simplified "equivalent" normal form for OBMDPs (Proposition 2.3 later on shows that OBMDPs can always be translated efficiently into this normal form). We extend the notation for rules in the model to adopt actions and not only probabilities, i.e., we will be using $T_i \xrightarrow{a} T_j$, where $a \in \Gamma^i$, to denote a rule where a non-terminal T_i generates as a child (under player's choice of action $a \in \Gamma^i$) a copy of non-terminal T_j (with probability 1).

Definition 9. *An OBMDP is in* ***simple normal form (SNF)*** *if each non-terminal T_i is in one of three possible forms:*

- L-FORM*: T_i is a "probabilistic" (or "linear") non-terminal (i.e., the player has no choice of actions), and the associated rules for T_i are given by: $T_i \xrightarrow{p_{i,0}} \varnothing, T_i \xrightarrow{p_{i,1}} T_1, \ldots, T_i \xrightarrow{p_{i,n}} T_n$, where for all $0 \leq j \leq n$, $p_{i,j} \geq 0$ denotes the probability of each rule, and $\sum_{j=0}^{n} p_{i,j} = 1$.*

- Q-FORM*: T_i is a "branching" (or "quadratic") non-terminal, with a single associated rule (and no associated actions) of the form $T_i \xrightarrow{1} T_j \, T_r$.*

- M-FORM*: T_i is a "controlled" non-terminal, with a non-empty set of associated actions $\Gamma^i = \{a_1, \ldots, a_{m_i}\} \subseteq \Sigma$, and the associated rules have the form $T_i \xrightarrow{a_1} T_{j_1}, \ldots, T_i \xrightarrow{a_{m_i}} T_{j_{m_i}}$.*[4]

A *derivation* for an OBMDP, starting at some start non-terminal $T_{start} \in V$, is a (possibly infinite) labeled ordered tree, $X = (B, s)$, defined as follows. The set of nodes $B \subseteq \{l, r, u\}^*$ of the tree, X, is a *prefix-closed* subset of $\{l, r, u\}^*$.[5] So each node in B is a string over $\{l, r, u\}$, and if $w = w'a \in B$, where $a \in \{l, r, u\}$, then $w' \in B$. As usual, when $w \in B$ and $w' = wa \in B$, for some $a \in \{l, r, u\}$, we call w the *parent* of w', and we call w' a *child* of w in the tree. A *leaf* of B is a node $w \in B$ that has no children in B. Let $\mathcal{L}_B \subseteq B$ denote the set of all leaves in B. The *root* node is the empty string e (note that B is prefix-closed, so $e \in B$). The function $s : B \to V \cup \{\varnothing\}$ assigns either a non-terminal or the empty symbol as a label to each node of the tree, and must satisfy

[4] We assume, without loss of generality, that for $0 \leq t < t' \leq m_i$, $T_{j_t} \neq T_{j_{t'}}$.

[5] Here 'l', 'r', and 'u', stand for 'left', 'right', and 'unique' child, respectively.

the following conditions: Firstly, $s(e) = T_{start}$, in other words the root must be labeled by the start non-terminal; Inductively, if for any *non-leaf* node $w \in B \setminus \mathcal{L}_B$ we have $s(w) = T_i$, for some $T_i \in V$, then:

- if T_i is a Q-form (branching) non-terminal, whose associated unique rule is $T_i \xrightarrow{1} T_j\ T_{j'}$, then w must have exactly two children in B, namely $wl \in B$ and $wr \in B$, and moreover we must have $s(wl) = T_j$ and $s(wr) = T_{j'}$.

- if T_i is a L-form (probabilistic) non-terminal, then w must have exactly one child in B, namely wu, and it must be the case that either $s(wu) = T_j$, where there exists some rule $T_i \xrightarrow{p_{i,j}} T_j$ with a positive probability $p_{i,j} > 0$, or else $s(wu) = \varnothing$, where there exists a rule $T_i \xrightarrow{p_{i,0}} \varnothing$, with an empty right-hand side, and a positive probability $p_{i,0} > 0$.

- if T_i is a M-form (controlled) non-terminal, then w must have exactly one child in B, namely wu, and it must be the case that $s(wu) = T_{j_z}$, where there exists some rule $T_i \xrightarrow{a_z} T_{j_z}$, associated with some action $a_z \in \Gamma^i$, having non-terminal T_i as its left-hand side.

A derivation $X = (B, s)$ is *finite* if the set B is finite. A derivation $X' = (B', s')$ is called a *subderivation* of a derivation $X = (B, s)$, if $B' \subseteq B$ and $s' = s|_{B'}$ (i.e., s' is the function s, restricted to the domain B'). We use $X' \preceq X$ to denote the fact that X' is a subderivation of X.

A *complete* derivation, or a *play*, $X = (B, s)$, is by definition a derivation in which for all leaves $w \in \mathcal{L}_B$, $s(w) = \varnothing$. For a play $X = (B, s)$, and a node $w \in B$, we define the *subplay of X rooted at w*, to be the play $X^w = (B^w, s^w)$, where $B^w = \{w' \in \{l, r, u\}^* \mid ww' \in B\}$ and $s^w : B^w \to V \cup \{\varnothing\}$ is given by, $s^w(w') := s(ww')$ for all $w' \in B^w$.[6]

Consider any derivation $X = (B, s)$, and any node $w = w_1 \dots w_m \in B$, where $w_t \in \{l, r, u\}$ for all $t \in [m]$. We define the *ancestor history* of w to be a sequence $h_w \in V(\{l, r, u\} \times V)^*$, given by $h_w := s(e)(w_1, s(w_1))(w_2, s(w_1 w_2))(w_3, s(w_1 w_2 w_3)) \dots (w_m, s(w_1 w_2 \dots w_m))$. In other words, the ancestor history h_w of node w specifies the sequence of moves that determines each ancestor of w (starting at root node e and including w itself), and also specifies the sequence of non-terminals that label each ancestor of w.

[6]To avoid confusion, note that subderivation and subplay have very different meanings. Saying derivation X is a "subderivation" of derivation X', means that in a sense X is a "prefix" of X', as an ordered tree. Saying play X is a subplay of play X', means X is a "suffix" of X', more specifically X is a subtree rooted at a specific node of X'.

For an OBMDP, $\mathcal{A}$, a sequence $h \in V(\{l,r,u\} \times V)^*$ is called a *valid* ancestor history if there is some derivation $X = (B',s')$ of $\mathcal{A}$, and node $w \in B'$ such that $h = h_w$. We define the *current non-terminal* of such a valid ancestor history h to be $s'(w)$. In other words, it is the non-terminal that labels the last node of the ancestor history h. Let $\texttt{current}(h)$ denote the current non-terminal of h. Let $H_{\mathcal{A}} \subseteq V(\{l,r,u\} \times V)^*$ denote the set of all valid ancestor histories of $\mathcal{A}$. A valid ancestor history $h \in H_{\mathcal{A}}$ is said to *belong to the controller*, if $\texttt{current}(h)$ is a M-form (controlled) non-terminal. Let $H^C_{\mathcal{A}}$ denote the set of all valid ancestor histories of the OBMDP, $\mathcal{A}$, that belong to the controller.

For an OBMDP, $\mathcal{A}$, a general *history* of the OBMDP process at time $t \in \mathbb{N}$ is a finite derivation (i.e., a finite labeled ordered tree) of depth t, which also contains the information regarding all the past actions and rules applied up to the "generation" of t. A general *strategy* for the controller is a function that maps every finite derivation, $X = (B,s)$, to a function that maps each leaf $w \in \mathcal{L}_B$, such that $s(w) \neq \varnothing$ and $s(w)$ is a M-form non-terminal, to a probability distribution on the actions Γ^i, assuming $s(w) = T_i$. Note that the strategy can choose different distributions on actions at different occurrences of the same non-terminal in the derivation tree, even when these occurrences happen to be "siblings" in the tree. We also define another (restricted) notion of a strategy, utilizing the (weaker) notion of an ancestor history. An *ancestral strategy* for the controller is a function, $\sigma : H^C_{\mathcal{A}} \to \Delta(\Sigma)$, from the set of valid ancestor histories belonging to the controller, to probability distributions on actions, such that moreover for any $h \in H^C_{\mathcal{A}}$, if $\texttt{current}(h) = T_i$, then $\sigma(h) \in \Delta(\Gamma^i)$. (In other words, the probability distribution must have support only on the actions available at the current non-terminal.)

Now is the moment to clarify something important. There is a reason why we have the restricted definition of an ancestral strategy in the context of OBMDPs. As later discussed in subsection 2.4.1, computing the optimal (single-target) reachability probabilities in OBMDPs is equivalent to computing the optimal (single-target) reachability probabilities in BMDPs. Same holds for the objective of extinction (called termination in the model of OBMDPs). These equivalences hold also under the restriction to ancestral strategies in the context of OBMDPs (the reasons will become clear in subsection 2.4.1). Furthermore, we point out that even under the stronger and more general notion of a strategy, where the history is the entire finite tree up to the current generation, it was shown in [ESY18, Example 3.2] that there may be no optimal strategy for the player maximizing the reachability probability.

Later in Section 5.1, we pose some open problems related to the general notion of a strategy in the context of OBMDPs. But for now we make the following restriction for OBMDPs.

From now on in the rest of the thesis, every mention of a "strategy" in the context of OBMDPs will refer to the notion of an ancestral strategy. That is, for the rest of thesis, for OBMDPs we restrict ourselves to ancestral strategies. *By contrast, in the context of BMDPs, BSSGs and BCSGs, we assume that "strategies" have as a history the entire finite tree up to the "current generation", as defined in Section 2.2.*

For an OBMDP, $\mathcal{A}$, let Ψ be the set of all *strategies* (we now mean ancestral strategies). We say $\sigma \in \Psi$ is *deterministic* if for all $h \in H^C_{\mathcal{A}}$, $\sigma(h)$ puts probability 1 on a single action. We say $\sigma \in \Psi$ is *static* if for each M-form (controlled) non-terminal T_i, there is some distribution $\delta_i \in \Delta(\Gamma^i)$, such that for any $h \in H^C_{\mathcal{A}}$ with $\texttt{current}(h) = T_i$, $\sigma(h) = \delta_i$. In other words, a static strategy σ plays, for each M-form non-terminal T_i, exactly the same distribution on actions at every occurrence of T_i in the tree (play), regardless of the ancestor history.

For an OBMDP, $\mathcal{A}$, fixing a start non-terminal T_i, and fixing a strategy σ for the controller, determines a stochastic process that generates a random play, as follows. The process generates a sequence of finite derivations, $X_0, X_1, X_2, X_3, \ldots$, one for each "generation", such that for all $t \in \mathbb{N}$, $X_t \preceq X_{t+1}$. $X_0 = (B_0, s_0)$ is the initial derivation, at generation 0, and consists of a single (root) node $B_0 = \{e\}$, labeled by the start non-terminal, $s_0(e) = T_i$.[7] Inductively, for all $t \in \mathbb{N}$ the derivation $X_{t+1} = (B_{t+1}, s_{t+1})$ is obtained from $X_t = (B_t, s_t)$ as follows. For each leaf $w \in \mathcal{L}_{B_t}$:

- if $s_t(w) = T_i$ is a Q-form (branching) non-terminal, whose associated unique rule is $T_i \xrightarrow{1} T_j\, T_{j'}$, then w must have exactly two children in B_{t+1}, namely $wl \in B_{t+1}$ and $wr \in B_{t+1}$, and moreover we must have $s_{t+1}(wl) = T_j$ and $s_{t+1}(wr) = T_{j'}$.

- if $s_t(w) = T_i$ is a L-form (probabilistic) non-terminal, then w has exactly one child in B_{t+1}, namely wu, and for each rule $T_i \xrightarrow{p_{i,j}} T_j$ with $p_{i,j} > 0$, the probability that $s_{t+1}(wu) = T_j$ is $p_{i,j}$, and likewise when $T_i \xrightarrow{p_{i,0}} \varnothing$ is a rule with $p_{i,0} > 0$, then $s_{t+1}(wu) = \varnothing$ with probability $p_{i,0}$.

[7] We can assume, without loss of generality, that the initial generation consists of a single given root non-terminal, because for any given collection $\mu \in V^*$ of multiple roots, we can always add an auxiliary non-terminal T_f to the original OBMDP, where $\Gamma^f = \{a\}$ and the set $R(T_f, a)$ contains a single probabilistic rule, $T_f \xrightarrow{1} \mu$.

- if $s_t(w) = T_i$ is a M-form (controlled) non-terminal, then w has exactly one child in B_{t+1}, namely wu, and for each action $a_z \in \Gamma^i$, with probability $\sigma(h_w)(a_z)$, $s_{t+1}(wu) = T_{j_z}$, where $T_i \xrightarrow{a_z} T_{j_z}$ is the rule associated with a_z.

There are no other nodes in B_{t+1}. In particular, if $s_t(w) = \varnothing$, then in B_{t+1} the node w has no children. This defines a stochastic process, $X_0, X_1, X_2, \ldots$, where $X_t \preceq X_{t+1}$, for all $t \in \mathbb{N}$, and such that there is a unique play, $X = \lim_{t\to\infty} X_t$, such that $X_t \preceq X$ for all $t \in \mathbb{N}$.

In this sense, the random process defines a probability space of plays. To be more precise, let $\mathcal{A}$ be an OBP and, for any finite derivation (tree) X, let $C_{\mathcal{A}}(X) := \{X' \mid X'$ is a derivation and $X \preceq X'\}$ be the *cylinder* over X, i.e., $C_{\mathcal{A}}(X)$ is the set of derivations or plays X' such that X is a subderivation of X'. Then $\mathcal{A}$ defines the probability space $(\Omega, \mathbb{F}, \mathbb{P})$, where the sample space Ω is the set of plays. The σ-algebra, $\mathbb{F} \subseteq 2^{\Omega}$, of measurable events associated with plays of OBP, $\mathcal{A}$, is the σ-algebra generated by the cylinders $\{C_{\mathcal{A}}(X) \mid X$ is a finite derivation$\}$. The probability measure, $\mathbb{P} : \mathbb{F} \to [0,1]$, is the uniquely determined measure by specifying the probabilities of the cylinders, where each such probability, $\mathbb{P}[C_{\mathcal{A}}(X)]$, is simply the product of all rules in the finite derivation X.

For our purposes, an *objective* is specified by a property (i.e., a measurable set), $\mathcal{F}$, of plays, whose probability the player wishes to optimize (maximize or minimize). Different objectives can be considered for OBMDPs (and for OBPs and their game extensions). Section 2.2 defined the objectives of *termination* (called *extinction* for BPs) and *(single-target) reachability*, and in subsection 2.4.1 we will observe that in fact under both objectives the models of BPs and OBPs are equivalent. In the termination objective, the aim of the player is to optimize (maximize or minimize) the probability that the process terminates, i.e., that the generated play is finite; and in the (single-target) reachability objective, the goal of the player is to optimize (maximize or minimize) the probability of the play containing a given target non-terminal.

Chapter 4 analyses *multi-objective reachability*, which is a natural extension of the previously studied (single-target) reachability. In the multi-objective setting:

- we have multiple given target non-terminals, and the goal is to optimize each of the respective probabilities of achieving multiple given objectives, each one being a Boolean combination (using union and intersection) of reachability and non-reachability properties over different target non-terminals. Of course, there may be tradeoffs between optimizing the probabilities of achieving the different objectives.

To formalize things, we need some notation. Given a target non-terminal T_q, $q \in [n]$, let $Reach(T_q)$ denote the set of plays that contain some copy (some node) of non-terminal T_q. Respectively, let $Reach^{\complement}(T_q)$ denote the complement event, i.e., the set of plays that do *not* contain a node labelled by non-terminal T_q. For any measurable set (i.e., property) of plays, $\mathcal{F}$, and for any strategy σ for the player and a given starting non-terminal T_i, we denote by $Pr^{\sigma}_{T_i}[\mathcal{F}]$ the probability that, starting at a non-terminal T_i and under strategy σ, the generated play is in the set $\mathcal{F}$. Let $Pr^{*}_{T_i}[\mathcal{F}] := \sup_{\sigma\in\Psi} Pr^{\sigma}_{T_i}[\mathcal{F}]$.

The *quantitative multi-objective decision* problem for OBMDPs is the following problem. We are given an OBMDP, a starting non-terminal $T_s \in V$, a collection of objectives (properties) $\mathcal{F}_1, \ldots, \mathcal{F}_k$ and corresponding probabilities $p_1, \ldots, p_k$. The problem asks to decide whether there exists a strategy $\sigma' \in \Psi$ such that $\bigwedge_{i\in[k]} Pr^{\sigma'}_{T_s}[\mathcal{F}_i]\triangle_i p_i$ holds, where $\triangle_i \in \{<, \leq, =, \geq, >\}$. Observe that the clauses (i.e., the probability queries $Pr^{\sigma'}_{T_s}[\mathcal{F}_i]\triangle_i p_i$, for any $i \in [k]$) with $\triangle_i = \leq$ and $\triangle_i = \geq$ inequalities can be converted to ask whether either $Pr^{\sigma'}_{T_s}[\mathcal{F}_i] = p_i$, or $Pr^{\sigma'}_{T_s}[\mathcal{F}_i] < p_i$ (respectively, $Pr^{\sigma'}_{T_s}[\mathcal{F}_i] > p_i$). Moreover, we could in general allow for any Boolean combination of clauses (not just a conjunction). In any case, the whole query can be put into a disjunctive normal form and the quantification over strategies can be pushed inside the disjunction. So any multi-objective query can eventually be transformed into a disjunction of finite number of (smaller) queries. (Note that, of course, this number can be exponential in the size of the original multi-objective query.) Hence, we can define a multi-objective decision problem only as a conjunction of equality and strict inequality queries.

One could also ask the *limit* version of this question. For instance, whether for all $\varepsilon > 0$, there exists a strategy $\sigma'_\varepsilon \in \Psi$, such that $\bigwedge_{i\in[k]} Pr^{\sigma'_\varepsilon}_{T_s}[\mathcal{F}_i] \geq p_i - \varepsilon$. Moreover, we can also ask quantitative questions regarding computing (or approximating) the *Pareto curve* for the multiple objectives, but we will not consider such questions in this thesis (Section 5.1 leaves such questions as future work).

The *qualitative almost-sure multi-objective decision* problem for OBMDPs is the special case where $p_i = \{0, 1\}$ for each $i \in [k]$. In other words, this problem is phrased as asking whether, starting at a given non-terminal $T_s \in V$, there exists a strategy $\sigma \in \Psi$ such that $\bigwedge_{i\in[k]} Pr^{\sigma}_{T_s}[\mathcal{F}_i]\triangle_i\{0, 1\}$ (where as mentioned $\triangle_i \in \{<, =, >\}$). We can simplify the expression by transforming clauses of the form $Pr^{\sigma}_{T_s}[\mathcal{F}_i] > 0$ and $Pr^{\sigma}_{T_s}[\mathcal{F}_i] = 0$ into $Pr^{\sigma}_{T_s}[\mathcal{F}_i^{\complement}] < 1$ and $Pr^{\sigma}_{T_s}[\mathcal{F}_i^{\complement}] = 1$, respectively, where each $\mathcal{F}_i^{\complement}$ is the complement objective of $\mathcal{F}_i$.

Then, for a strategy $\sigma \in \Psi$ and a starting non-terminal $T_s \in V$, the expression can be rephrased as: $\bigwedge_{i\in[k_1]} Pr^{\sigma}_{T_s}[\mathcal{F}_i] < 1 \wedge \bigwedge_{i\in[k_2]} Pr^{\sigma}_{T_s}[\mathcal{F}_i] = 1$, where $k_1 + k_2 = k$. And by Proposition 2.2(1.) below, the qualitative (almost-sure) multi-objective decision problem reduces to asking whether there exists a strategy $\sigma' \in \Psi$ such that $\bigwedge_{i\in[k_1]} Pr^{\sigma'}_{T_s}[\mathcal{F}_i] < 1 \wedge Pr^{\sigma'}_{T_s}[\bigcap_{i\in[k_2]} \mathcal{F}_i] = 1$.

The *qualitative limit-sure multi-objective decision* problem for OBMDPs asks to decide whether, for every $\varepsilon > 0$, there exists a strategy $\sigma'_\varepsilon \in \Psi$ for the player such that $\bigwedge_{i\in[k]} Pr^{\sigma'_\varepsilon}_{T_s}[\mathcal{F}_i] \geq 1-\varepsilon$. Again by Proposition 2.2(5.) below, it follows that the qualitative limit-sure multi-objective decision problem can be rephrased as asking whether, for all $\varepsilon > 0$, there exists a strategy $\sigma'_\varepsilon \in \Psi$ such that $Pr^{\sigma'_\varepsilon}_{T_s}[\bigcap_{i\in[k]} \mathcal{F}_i] \geq 1-\varepsilon$.

The following Proposition shows scenarios where the qualitative multi-objective problem for OBMDPs can be rephrased as a qualitative single-objective problem, where the single objective is a Boolean combination of the given multiple objectives.

Proposition 2.2. *Given an OBMDP, with a starting non-terminal $T_s \in V$ and a collection $\mathcal{F}_1,\ldots,\mathcal{F}_k$ of k objectives:*

(1.) $\exists\sigma' \in \Psi: \bigwedge_{i\in[k]} Pr^{\sigma'}_{T_s}[\mathcal{F}_i] = 1$ if and only if $\exists\sigma' \in \Psi: Pr^{\sigma'}_{T_s}[\bigcap_{i\in[k]} \mathcal{F}_i] = 1$.

(2.) $\exists\sigma' \in \Psi: \bigvee_{i\in[k]} Pr^{\sigma'}_{T_s}[\mathcal{F}_i] < 1$ if and only if $\exists\sigma' \in \Psi: Pr^{\sigma'}_{T_s}[\bigcap_{i\in[k]} \mathcal{F}_i] < 1$.

(3.) $\exists\sigma' \in \Psi: \bigwedge_{i\in[k]} Pr^{\sigma'}_{T_s}[\mathcal{F}_i] = 0$ if and only if $\exists\sigma' \in \Psi: Pr^{\sigma'}_{T_s}[\bigcup_{i\in[k]} \mathcal{F}_i] = 0$.

(4.) $\exists\sigma' \in \Psi: \bigvee_{i\in[k]} Pr^{\sigma'}_{T_s}[\mathcal{F}_i] > 0$ if and only if $\exists\sigma' \in \Psi: Pr^{\sigma'}_{T_s}[\bigcup_{i\in[k]} \mathcal{F}_i] > 0$.

Moreover, in each of the equivalence statements (1.) - (4.), a witness strategy σ' for one of the sides is also a witness strategy for the other.

(5.) Similar equivalence holds for the qualitative limit-sure multi-objective problem: $\forall\varepsilon > 0, \exists\sigma'_\varepsilon \in \Psi: \bigwedge_{i\in[k]} Pr^{\sigma'_\varepsilon}_{T_s}[\mathcal{F}_i] \geq 1-\varepsilon$ if and only if $\forall\varepsilon > 0, \exists\sigma'_\varepsilon \in \Psi: Pr^{\sigma'_\varepsilon}_{T_s}[\bigcap_{i\in[k]} \mathcal{F}_i] \geq 1-\varepsilon$.

And from a witness strategy σ'_ε (for $\varepsilon > 0$) for one of the two sides a witness strategy $\sigma''_{\varepsilon'}$ (for a potentially different $\varepsilon' > 0$) can be obtained for the other.

Proof.

(1.). For one direction of the statement, suppose there is a strategy $\sigma' \in \Psi$ for the player such that $Pr^{\sigma'}_{T_s}[\bigcap_{i\in[k]} \mathcal{F}_i] = 1$, i.e., almost-surely all objectives are satisfied in the same generated play. It follows that $Pr^{\sigma'}_{T_s}[\bigcup_{i\in[k]} \mathcal{F}_i^{\complement}] = 0$. Clearly, for each $i \in [k]$, $Pr^{\sigma'}_{T_s}[\mathcal{F}_i^{\complement}] = 0$ and hence, for each $i \in [k]: Pr^{\sigma'}_{T_s}[\mathcal{F}_i] = 1$.

Showing the other direction, suppose that there exists a strategy $\sigma' \in \Psi$ for the player such that $\bigwedge_{i\in[k]} Pr^{\sigma'}_{T_s}[\mathcal{F}_i] = 1$. Then, $\forall i \in [k]$, $Pr^{\sigma'}_{T_s}[\mathcal{F}_i^{\complement}] = 0$. By the union bound, $Pr^{\sigma'}_{T_s}[\bigcup_{i\in[k]} \mathcal{F}_i^{\complement}] = 0$ and, hence, $Pr^{\sigma'}_{T_s}[\bigcap_{i\in[k]} \mathcal{F}_i] = 1$.

(2.). For one direction of the statement, suppose there is a strategy $\sigma' \in \Psi$ such that $Pr^{\sigma'}_{T_s}[\bigcap_{i\in[k]} \mathcal{F}_i] < 1$. Then $Pr^{\sigma'}_{T_s}[\bigcup_{i\in[k]} \mathcal{F}_i^{\complement}] > 0$. Clearly, $\exists i' \in [k]$ such that $Pr^{\sigma'}_{T_s}[\mathcal{F}_{i'}^{\complement}] > 0$ (otherwise, by the union bound the probability of the union of the events is 0). Hence, $\bigvee_{i\in[k]} Pr^{\sigma'}_{T_s}[\mathcal{F}_i] < 1$.

As for the other direction, suppose there is a strategy $\sigma' \in \Psi$ and some $i' \in [k]$ such that $Pr^{\sigma'}_{T_s}[\mathcal{F}_{i'}] < 1$. Then $Pr^{\sigma'}_{T_s}[\bigcap_{i\in[k]} \mathcal{F}_i] \leq Pr^{\sigma'}_{T_s}[\mathcal{F}_{i'}] < 1$.

(3.) and (4.) follow directly from (1.) and (2.), respectively.

(5.). For one direction of the statement, suppose that for every $\varepsilon > 0$ there is a strategy $\sigma'_\varepsilon \in \Psi$ such that $Pr^{\sigma'_\varepsilon}_{T_s}[\bigcap_{i\in[k]} \mathcal{F}_i] \geq 1-\varepsilon$, i.e., limit-surely (with probability arbitrarily close to 1) all objectives are satisfied in the same generated play. It follows that $Pr^{\sigma'_\varepsilon}_{T_s}[\bigcup_{i\in[k]} \mathcal{F}_i^{\complement}] \leq \varepsilon$. Clearly, for each $i \in [k]$, $Pr^{\sigma'_\varepsilon}_{T_s}[\mathcal{F}_i^{\complement}] \leq \varepsilon$, and hence, for each $i \in [k] : Pr^{\sigma'_\varepsilon}_{T_s}[\mathcal{F}_i] \geq 1-\varepsilon$.

Showing the other direction, suppose that for every $\varepsilon > 0$ there exists a strategy $\sigma'_\varepsilon \in \Psi$ such that $\bigwedge_{i\in[k]} Pr^{\sigma'_\varepsilon}_{T_s}[\mathcal{F}_i] \geq 1-\varepsilon$. Then, for every $i \in [k]$, $Pr^{\sigma'_\varepsilon}_{T_s}[\mathcal{F}_i^{\complement}] \leq \varepsilon$. By the union bound, $Pr^{\sigma'_\varepsilon}_{T_s}[\bigcup_{i\in[k]} \mathcal{F}_i^{\complement}] \leq k\varepsilon$, and hence, $Pr^{\sigma'_\varepsilon}_{T_s}[\bigcap_{i\in[k]} \mathcal{F}_i] \geq 1-k\varepsilon$. So for any $\varepsilon > 0$, let $\varepsilon' := \varepsilon/k$ and $\sigma_\varepsilon := \sigma'_{\varepsilon'}$, where $\sigma'_{\varepsilon'}$ satisfies $\bigwedge_{i\in[k]} Pr^{\sigma'_{\varepsilon'}}_{T_s}[\mathcal{F}_i] \geq 1-\varepsilon' = 1-\varepsilon/k$. Then it follows that $Pr^{\sigma_\varepsilon}_{T_s}[\bigcap_{i\in[k]} \mathcal{F}_i] \geq 1-k\varepsilon' = 1-\varepsilon$. □

In Chapter 4, we address certain cases of the qualitative multi-objective reachability decision problem for OBMDPs. We are given a collection of *generalized* reachability objectives $\mathcal{F}_1, \dots, \mathcal{F}_k$, where each such generalized reachability objective $\mathcal{F}_i$, $i \in [k]$, represents a set of plays described by a Boolean combination over the sets (of plays) $Reach(T_q)$, $T_q \in V$, using the set operations *union*, *intersection* and *complementation*. That is, each generalized reachability objective $\mathcal{F}_i$, $i \in [k]$, is of the form $\bigcap_{t\in[z_i]}(\bigcup_{t'\in[z_{i,t}]} \Phi(T_{q_{i,t,t'}}))$, where $\Phi \in \{Reach, Reach^{\complement}\}$, $T_{q_{i,t,t'}} \in V$ and the values $z_i, z_{i,t}$ are part of the objective $\mathcal{F}_i$.

We will show that, even in the case of having a single objective that asks to reach multiple target non-terminals from a given set in the same play, the almost-sure and limit-sure questions do not coincide and we give separate algorithms for detecting almost-sure and limit-sure multi-target reachability. (As later explained in Section

2.6, by combining results [ESY18, Theorems 9.3,9.4] for BMDPs and Proposition 2.4, then in the case of single-target reachability the almost-sure and limit-sure questions for OBMDPs do coincide.)

The following example indeed illustrates that there are OBMDPs where, even though the supremum probability of reaching all target non-terminals from a given set in the same play is 1, there may not exist a strategy for the player that actually achieves probability exactly 1. Recall that we have already restricted ourselves to ancestral strategies in the context of OBMDPs.

Example 2.1. *(The qualitative almost-sure and limit-sure multi-target reachability problems for OBMDPs do not coincide.) Consider the following OBMDP with non-terminals set $\{M, A, R_1, R_2\}$, where R_1 and R_2 are the target non-terminals. M is the only "controlled" non-terminal, and the rules are:*

$$M \xrightarrow{a} M\,A \qquad\qquad A \xrightarrow{1/2} R_1$$

$$M \xrightarrow{b} R_2 \qquad\qquad A \xrightarrow{1/2} \varnothing$$

The supremum probability, $Pr^*_M[Reach(R_1) \cap Reach(R_2)]$, starting at a non-terminal M, of reaching both targets is 1. To see this, for any $\varepsilon > 0$, let the strategy keep choosing deterministically action a until $l := \lceil \log_2(\frac{1}{\varepsilon}) \rceil$ copies of non-terminal A have been created, i.e., until the play reaches generation l. Then in the (unique) copy of non-terminal M in generation l the strategy switches deterministically to action b. The probability of reaching target R_2 is 1. The probability of reaching target R_1 is $1 - 2^{-l} \geq 1 - \varepsilon$. The player can delay arbitrarily long the moment when to switch from choosing action a to choosing action b for a non-terminal M. Hence, $Pr^*_M[Reach(R_1) \cap Reach(R_2)] = 1$.

However, $\nexists \sigma \in \Psi : Pr^{\sigma}_M[Reach(R_1) \cap Reach(R_2)] = 1$. To see this, note that if the strategy ever puts a positive probability on action b in any "round", then with a positive probability target R_1 will not be reached in the play. So, to reach target R_1 with probability 1, the strategy must deterministically choose action a forever, from every occurrence of non-terminal M. But if it does this, the probability of reaching target R_2 would be 0. □

The following Proposition is easy to prove (similar to Proposition 2.1 and [ESY18, Proposition 2.6]) and shows that we can always efficiently convert an OBMDP into its SNF form (Definition 9).

Proposition 2.3. *Every OBMDP, $\mathcal{A}$, can be converted in P-time to an "equivalent" OBMDP, $\mathcal{A}'$, in SNF form, such that $|\mathcal{A}'| \in O(|\mathcal{A}|)$. More precisely, the non-terminals $V = \{T_i \mid i \in [n]\}$ of $\mathcal{A}$ are a subset of the non-terminals of $\mathcal{A}'$, and any strategy σ of $\mathcal{A}$ can be converted to a strategy σ' of $\mathcal{A}'$ (and vice versa), such that starting at any non-terminal $T_s \in V$, and for any generalized reachability objective $\mathcal{F}$ (with respect to the non-terminals in $\mathcal{A}$), using the strategies σ and σ' in $\mathcal{A}$ and $\mathcal{A}'$, respectively, the probability that the resulting play is in the set of plays, $\mathcal{F}$, is the same in both $\mathcal{A}$ and $\mathcal{A}'$.*

Proof. For a rule $T_i \xrightarrow{p_r} s_r$, $s_r \in V^*$, in $\mathcal{A}$ and a non-terminal $T_j \in V$, let $m_{r,j} := |\{d \mid (s_r)_d = T_j,\ 1 \leq d \leq |s_r|\}|$ be the number of copies of T_j in string s_r. We use the following procedure to convert, in P-time, any OBMDP, $\mathcal{A}$, into its SNF-form OBMDP, $\mathcal{A}'$.

1. Initialize $\mathcal{A}'$ by adding all the non-terminals $T_i \in V$ from $\mathcal{A}$ and their corresponding action sets Γ^i.

2. For each non-terminal T_i, such that $m_{r,i} > 1$ for some non-terminal T_j, action $a \in \Gamma^j$ and rule $r \in R(T_j,a)$ from $\mathcal{A}$, create new non-terminals $T_{i_1},\ldots,T_{i_z}$ in $\mathcal{A}'$ where $z = \lfloor \log_2(\max_{r' \in R}\{m_{r',i}\}) \rfloor$. Then add the rules $T_{i_1} \xrightarrow{1} T_i\ T_i$, $T_{i_2} \xrightarrow{1} T_{i_1}\ T_{i_1}$, $\ldots$, $T_{i_z} \xrightarrow{1} T_{i_{z-1}}\ T_{i_{z-1}}$ to $\mathcal{A}'$. For every rule r in OBMDP, $\mathcal{A}$, where $m_{r,i} > 1$, if the binary representation of $m_{r,i}$ is $l_z \ldots l_2 l_1 l_0$, then we remove all copies of T_i in string s_r (i.e., the right-hand side of rule r) and add a copy of non-terminal T_{i_t} to string s_r if bit $l_t = 1$, for every $0 \leq t \leq z$. After this step, for every rule r, the string s_r consists of at most one copy of any non-terminal.

3. For each non-terminal T_i, for each action $a_d \in \Gamma^i$, create a new non-terminal T_d in $\mathcal{A}'$ and add the rule $T_i \xrightarrow{a_d} T_d$ to $\mathcal{A}'$.

4. Next, for each such new non-terminal T_d from point 3., for each rule r from set $R(T_i,a_d)$ in $\mathcal{A}$: if $s_r = \varnothing$ (i.e., the set of children under rule r is empty), then add the rule $T_d \xrightarrow{p_r} \varnothing$ to $\mathcal{A}'$; if the set of children consists of a single copy of some non-terminal T_j, then add the rule $T_d \xrightarrow{p_r} T_j$ to $\mathcal{A}'$; and if the set of children is larger and s_r does not have an associated non-terminal already, then create a new non-terminal T_{d_r}, associated with string s_r, in $\mathcal{A}'$ and add the rule $T_d \xrightarrow{p_r} T_{d_r}$ to $\mathcal{A}'$.

5. Next, for each such new non-terminal T_{d_r}, associated with $s_r, r \in R(T_i,a_d)$, where s_r contains $m \geq 2$ non-terminals $T_{j_1},\ldots,T_{j_m}$: if $m = 2$, add rule $T_{d_r} \xrightarrow{1} T_{j_1}\ T_{j_2}$ to

$\mathcal{A}'$; and if $m > 2$, create $m-2$ new non-terminals $T_{l_1}, \ldots, T_{l_{m-2}}$ in $\mathcal{A}'$ and add the rules $T_{d_r} \xrightarrow{1} T_{j_1}\ T_{l_1},\ T_{l_1} \xrightarrow{1} T_{j_2}\ T_{l_2},\ T_{l_2} \xrightarrow{1} T_{j_3}\ T_{l_3},\ \ldots,\ T_{l_{m-2}} \xrightarrow{1} T_{j_{m-1}}\ T_{j_m}$ to $\mathcal{A}'$.

Now all non-terminals are of form L, Q or M.

The above procedure converts any OBMDP, $\mathcal{A}$, into one in SNF form by introducing $O(|\mathcal{A}|)$ new non-terminals and blowing up the size of $\mathcal{A}$ by a constant factor $O(1)$. Moreover, any strategy σ of the original OBMDP, $\mathcal{A}$, can be converted to a strategy σ' of the SNF-form OBMDP, $\mathcal{A}'$, (and vice versa) such that, under strategies σ and σ' in $\mathcal{A}$ and $\mathcal{A}'$, respectively, the probability that the resulting play is in the set of plays of a given generalized reachability objective $\mathcal{F}$ (over the non-terminals of $\mathcal{A}$) is the same in both $\mathcal{A}$ and $\mathcal{A}'$. □

From now on, throughout the rest of the thesis unless explicitly specified, *we may assume, without loss of generality, that any OBMDP is in SNF form.*

2.4.1 Equivalence between the models of BPs and OBPs

Recall the notation so far for the models of BPs and OBPs from Sections 2.2 and 2.4. In an OBMDP, $\mathcal{A}$, $Pr^{\sigma}_{T_i}[Reach^{\complement}(T_q)]$ denoted the probability in $\mathcal{A}$ of *not* reaching the given target non-terminal T_q, starting at a non-terminal T_i and under strategy $\sigma \in \Psi$ (where Ψ is the set of all strategies for $\mathcal{A}$). In an OBP, where there is only one trivial strategy, the same probability is simply denoted as $Pr_{T_i}[Reach^{\complement}(T_q)]$. Similarly, in a BP, $\mathcal{B}$, g^*_i denoted the non-reachability probability of the given target type T_q, starting at an object of type T_i. The following proposition shows that OBPs and BPs are equivalent with respect to the single-target (non-)reachability objective.

Proposition 2.4. *Every OBP, $\mathcal{A}$, can be translated in linear time to a BP, $\mathcal{B}$, such that there is a mapping from the non-terminals T_i in $\mathcal{A}$ to the types T_i in $\mathcal{B}$ such that, for a given target non-terminal (type) T_q, $Pr_{T_i}[Reach^{\complement}(T_q)] = g^*_i$.*

Conversely, every BP, $\mathcal{B}$, can be translated in polynomial time to an OBP, $\mathcal{A}$, such that there is a mapping from the types T_i in $\mathcal{B}$ to the non-terminals T_i in $\mathcal{A}$ such that, for a given target type (non-terminal) T_q, $g^*_i = Pr_{T_i}[Reach^{\complement}(T_q)]$.

Proof. Given an OBP, $\mathcal{A}$, in SNF form (as per Definition 9) one can construct a BP, $\mathcal{B}$, with the same (non-)reachability probabilities for corresponding start types in the following way. Let $\mathcal{B}$ have the same set of types as the set of non-terminals in $\mathcal{A}$. For every non-terminal T_i in $\mathcal{A}$, for the corresponding type T_i in $\mathcal{B}$, create a set $R(T_i)$ of rules (recall that a BP is a BCSG where both players have only one available action in each type) that consists of all rules in $\mathcal{A}$, where T_i is the left-hand side, and the right-hand side is converted from a string of non-terminals to a multi-set over types. The rules probabilities are kept consistent and if T_i is a M-form non-terminal in $\mathcal{A}$ then it surely has a singleton action set $\Gamma^i = \{a\}$ and has a single rule $T_i \xrightarrow{a} T_j$. In $\mathcal{B}$, this rule will have probability 1.

Now notice that for $\mathcal{A}$ one can build the following PPS (in SNF-form as per Definition 6), $x = P(x)$. For any L-form non-terminal T_i in $\mathcal{A}$ with rules $T_i \xrightarrow{p_{i,0}} \varnothing, T_i \xrightarrow{p_{i,1}} T_1, \ldots, T_i \xrightarrow{p_{i,n}} T_n$, there is a variable x_i and an equation $x_i = P_i(x) := p_{i,0} + \sum_{j=1}^{n} p_{i,j} x_j$. For any Q-form non-terminal T_i in $\mathcal{A}$ with a rule $T_i \xrightarrow{1} T_j\ T_r$, there is a variable x_i and an equation $x_i = P_i(x) := x_j x_r$. And for any M-form non-terminal T_i in $\mathcal{A}$ (which surely has $\Gamma^i = \{a\}$ and has a single rule $T_i \xrightarrow{a} T_j$ since $\mathcal{A}$ is an OBP), there is a variable x_i and an equation $x_i = P_i(x) := Val([x_j]) = x_j$. From $\mathcal{B}$ one can also build a SNF-form PPS and by obvious and trivial adjustments (e.g., substitutions and removing redundant variables) the two PPSs are the same. All proofs and results for BPs are applicable, *mutatis mutandis*, to OBPs with respect to the (non-)reachability objective. The greatest fixed point of the PPS also provides the non-reachability probabilities for the OBP.

In the opposite direction, given a BP, $\mathcal{B}$, first construct a PPS, $x = P(x)$, from $\mathcal{B}$ (as shown later in Section 3.1) and then convert it into SNF form in polynomial time (see Proposition 2.1). The previous paragraph showed how from any SNF-form OBP we can construct a SNF-form PPS whose greatest fixed point is the vector of non-reachability probabilities for the OBP with respect to the given target. Clearly, reversing the construction provides a corresponding OBP, $\mathcal{A}$, for the PPS, which is constructed from BP, $\mathcal{B}$. □

Proposition 2.4 can clearly be extended to the MDP and (concurrent) game generalizations of the BPs and OBPs models. There is one important note to make here.

A careful look at the constructed (ε-)optimal strategies in the results of [ESY18] implies that all the qualitative reachability results and the quantitative approximation reachability results for BMDPs from [ESY18] apply, *mutatis mutandis*, also for reachability in OBMDPs even under the restricted notion of ancestral strategies. In the context of BMDPs we need the more general notion of a strategy due to the lack of ordering among objects in a generation. What is more, in the qualitative almost-sure winning strategies there is a construct, called *queen-and-workers* (which is also utilised in the almost-sure winning strategies of BCSGs in Section 3.4), such that in BMDPs it can be implemented only with the use of the more general notion of strategies. However, in the context of OBMDPs, this queen-and-worker construct can be implemented even with the restricted notion of ancestral strategies, due to the fact that there is an ordering among the non-terminals.

And essentially, for single-target reachability, almost-sure and limit-sure reachability for OBMDPs coincide also under the restriction to ancestral strategies that we have defined in Section 2.4, i.e., where choices are based only on the ancestor history (with ordering information) of each node in the ordered tree. The qualitative reachability results for BSSGs and BCSGs from [ESY18] and [EMSY19] (Chapter 3) also apply for the game generalizations of OBMDPs under the restriction to ancestral strategies, again due to the specific nature of the constructed almost-sure winning, limit-sure winning and spoiling strategies for the two players. Hence, our restriction to only ancestral strategies in the context of OBMDPs is entirely justified. We come back later in Section 5.1 to the different notions of a strategy in order to pose some open problems.

Let us also observe an equivalence between BPs and OBPs with respect to the termination/extinction objective. Any OBP, $\mathcal{A}$, defines a global infinite-state Markov chain, $M^{\mathcal{A}} = (Q, \Delta)$, where the global states Q are finite labeled ordered trees (i.e., finite derivations) and a transition $(X, p_{X,X'}, X') \in \Delta$ exists for global states $X, X' \in Q$ if and only if $X \preceq X'$ and in fact there is a sequence of rules and actions, $\beta = \langle r_1, \ldots, r_z \rangle$, such that X' can be derived from X in one generation step using β (i.e., such that the "current" generations of X and X' are, respectively, $T_{i_1} T_{i_2} \ldots T_{i_z}$ and $s_{r_1} s_{r_2} \ldots s_{r_z}$, where for every $j \in [z]$: $s_{r_j} \in V^*$; if T_{i_j} is of form L or Q, then r_j is the rule $T_{i_j} \xrightarrow{p_{r_j}} s_{r_j}$; and if T_{i_j} is of M-form, then $r_j \in \Gamma^{i_j}$ is an action satisfying $T_{i_j} \xrightarrow{r_j} s_{r_j}$, and let $p_{r_j} := 1$). The probability of the transition is $p_{X,X'} := \prod_{j \in [z]} p_{r_j}$.

From the very similar definitions of global infinite-state MCs for the models of BPs and OBPs, one can observe that computing the termination probabilities in the OBPs model is equivalent to computing the extinction probabilities in the BPs model. That is, starting at a given non-terminal T_i, computing the probability of generating a finite play (i.e., the termination probability) in an OBP, $\mathcal{A}$, is equivalent to computing the probability of process becoming extinct (i.e., the extinction probability) in a corresponding BP, $\mathcal{B}$, starting in an object of the corresponding type T_i. The reason is that for the objective of termination the fact that there is an ordering of the non-terminals on the right-hand side of rules in $\mathcal{A}$ and an ordering of the non-terminals in each generation of a play in $\mathcal{A}$ is completely irrelevant. Furthermore, the definitions of the Markov chains $M^{\mathcal{A}}$ and $M^{\mathcal{B}}$ (i.e., the global denumerable MCs for $\mathcal{A}$ and $\mathcal{B}$, respectively) preserve the transition probabilities between the corresponding global states.

This equivalence with respect to the termination (extinction) objective also holds for the MDP and game generalizations of the two models. Thus, all the results for the extinction objective for BPs (and their 1- and 2-player-controlled extensions) apply, *mutatis mutandis*, to the termination objective for OBPs (and their 1- and 2-player-controlled extensions). Again, as noted earlier, we assumed that we restrict ourselves to ancestral strategies in the context of OBMDPs (or their game generalizations). And the aforementioned equivalence holds under this restriction. That is due to the important observation that, in 1- or 2-player-controlled (O)BPs under the extinction (termination) objective, strategies having access to a history, that includes the entire finite tree up to the current generation, is irrelevant to the goal of forcing an extinction (a termination) for the subtree of descendants of each object of the current generation. In other words, in each object of the current generation, it is irrelevant for the termination (extinction) objective to have information regarding what is happening in other parts of the tree.

2.5 Further Stochastic Models

This section discusses other classes of infinite-state stochastic processes that are related to (Ordered) Branching Processes and, in particular, we provide here definitions and background for Stochastic Context-Free Grammars and Recursive Markov chains.

2.5.1 Stochastic Context-Free Grammars

Definition 10. *A* ***Stochastic Context-Free Grammar (SCFG)****, $\mathcal{S}$, is a classic stochastic process, represented by a tuple $\mathcal{S} = (\Xi, V, R)$, where Ξ is a finite set of terminal symbols, $V = \{T_1, \ldots, T_n\}$ is a finite set of non-terminals and R is a finite set of probabilistic rules $T_i \xrightarrow{p_r} s_r$ such that $T_i \in V$, $p_r \in (0,1] \cap \mathbb{Q}$ is the rule probability (assumed to be a rational number for computational purposes) and $s_r \in (\Xi \cup V)^*$ is a (possibly empty) ordered string of terminals and non-terminals. For every non-terminal $T_i \in V$, $\sum_{(T_i \xrightarrow{p_r} s_r) \in R} p_r = 1$.*

Given a starting non-terminal, a possibly infinite (parse) tree (i.e., a derivation) is formed via one of the following forms of *rule-expansion*: *left-most* derivation, where at any step the left-most non-terminal in the string of terminals and non-terminals is chosen to be expanded by a probabilistically selected rule; similarly, *right-most* derivation prioritizes the right-most non-terminal; and *simultaneous* derivation, which simultaneously expands all the non-terminals in the current string of terminals and non-terminals. The derivation process is said to *terminate* once a string of only terminals is generated. For a string $s \in \Xi^*$ of terminals, the probability p_s of s being generated by the SCFG is defined as the sum of probabilities of all derivations (according to one of the rule-expansion forms) that terminate in string s, where the probability of each such derivation is the product of the probabilities of all the rules used in the parse tree of that derivation. For a given non-terminal T_i, the probability $p(T_i)$ of the language generated by the SCFG, starting at a non-terminal T_i (or equivalently, starting at a non-terminal T_i the probability of the SCFG stochastic process terminating), is the sum of probabilities p_s of all possible strings $s \in \Xi^*$ that can be derived from starting non-terminal T_i under the specified form of rule-expansion.

Context-Free Markov Decision Processes (CF-MDPs), Simple Stochastic Games (CF-SSGs) and Concurrent Stochastic Games (CF-CSGs) are the 1-player, 2-player turn-based and 2-player concurrent generalizations of SCFGs.

SCFGs with respect to computing the probability of the generated language, regardless of the rule-expansion form, are clearly equivalent to Ordered BPs with respect to computing the termination probability. This is due to the very similar definitions of the two stochastic models. Even though SCFGs contain terminal symbols and OBPs do not, in both stochastic models termination is essentially defined as generating a finite tree. Furthermore, computing the probability of the generated language in SCFGs is also equivalent to computing the extinction probability in BPs (see [EY09, Theorems 2.3, 2.4] or, equivalently, Theorems 2.5, 2.6). However, with respect to the (single-target) reachability objective, SCFGs without terminal symbols are equivalent to OBPs (and hence, equivalent to BPs by Proposition 2.4) only under the *simultaneous* derivation form of rule-expansion. Section 2.6 elaborates on the differences and similarities between all these models.

2.5.2 Recursive Markov models

Recursive Markov chains, introduced in [EY09], generalize Branching Processes, Ordered Branching Processes and Stochastic Context-Free Grammars.

Definition 11 (cf. [EY09])**.** *A* ***Recursive Markov Chain (RMC)****, $\mathcal{R}$, is a tuple $\mathcal{R} = (A_1, \ldots, A_n)$, where each component graph $A_i = (N_i, B_i, Y_i, En_i, Ex_i, \delta_i)$ consists of:*

- *A set N_i of nodes.*

- *A subset of entry nodes $En_i \subseteq N_i$, and a subset of exit nodes $Ex_i \subseteq N_i$.*

- *A set B_i of boxes, and a mapping $Y_i : B_i \to [n]$ that assigns to every box (the index of) one of the components, $A_1, \ldots, A_n$. To each box $b \in B_i$, we associate a set of call ports, $Call_b = \{(b, en) \mid en \in En_{Y_i(b)}\}$, corresponding to the entries of the corresponding component, and a set of return ports, $Return_b = \{(b, ex) \mid ex \in Ex_{Y_i(b)}\}$, corresponding to the exits of the corresponding component.*

- *A transition relation δ_i, where transitions are of the form $(u, p_{u,v}, v)$ where:*

  1. *u (the source) is either a non-exit node $u \in N_i - Ex_i$, or a return port $u = (b, ex)$ of some box $b \in B_i$,*
  2. *v (the destination) is either a non-entry node $v \in N_i - En_i$, or a call port $u = (b, en)$ of some box $b \in B_i$,*
  3. *$p_{u,v} \in \mathbb{R}_{>0}$ is the transition probability from u to v,*
  4. *For each u, $\sum_{\{v' \mid (u, p_{u,v'}, v') \in \delta_i\}} p_{u,v'} = 1$, unless u is a call port or an exit node, neither of which have outgoing transitions, in which case by default $\sum_{v'} p_{u,v'} = 0$.*

As in the other models in the thesis, all transition probabilities are assumed to be rational, for computational purposes, and the size of an instance $\mathcal{R}$ is measured by its bit encoding length. For a component, A_i, the set of all nodes, call ports and return ports in the component are collectively referred to as *vertices* and denoted by V_i. For a RMC, $\mathcal{R}$, let $N := \bigcup_{i \in [n]} N_i$ be the set of all nodes, $V := \bigcup_{i \in [n]} V_i$ be the set of all vertices, $B := \bigcup_{i \in [n]} B_i$ be the set of all boxes, $Y := \bigcup_{i \in [n]} Y_i$ be the mapping $Y : B \to [n]$ of all boxes to components, and $\delta = \bigcup_{i \in [n]} \delta_i$ be the set of all transitions.

Any RMC, $\mathcal{R}$, defines a global infinite-state Markov chain, $M^{\mathcal{R}} = (Q, \Delta)$, where the global states $Q \subseteq B^* \times V$ are pairs $\langle \beta, u \rangle$, where β is a (possibly empty) sequence of boxes and $u \in V$. Informally, if one thinks of the components as functions in a program and of the RMC as the call graph of the program, β represents the stack of recursive calls in an execution of the program. More formally, by [EY09]:

(1.) for every $u \in V$, $\langle e, u \rangle \in Q$, where e is the empty string.
(2.) if $\langle \beta, u \rangle \in Q$ and $(u, p_{u,v}, v) \in \delta$, then $\langle \beta, v \rangle \in Q$ and $(\langle \beta, u \rangle, p_{u,v}, \langle \beta, v \rangle) \in \Delta$.
(3.) if $\langle \beta, (b, en) \rangle \in Q$, where $(b, en) \in Call_b$, then $\langle \beta b, en \rangle \in Q$ and $(\langle \beta, (b, en) \rangle, 1, \langle \beta b, en \rangle) \in \Delta$.

(4.) if $\langle \beta b, ex \rangle \in Q$, where $(b, ex) \in Return_b$, then $\langle \beta, (b, ex) \rangle \in Q$ and $(\langle \beta b, ex \rangle, 1, \langle \beta, (b, ex) \rangle) \in \Delta$.

Point (1.) depicts all possible initial states of $M^{\mathcal{R}}$; point (2.) represents transitions in $M^{\mathcal{R}}$ that correspond to transitions within a single component in $\mathcal{R}$; point (3.) depicts transitions in $M^{\mathcal{R}}$ from call ports to their respective entries of the respective component, i.e., that correspond to recursive calls in $\mathcal{R}$; and point (4.) depicts transitions in $M^{\mathcal{R}}$ that correspond to exits from recursive calls in $\mathcal{R}$ and the process returning from the entered component to the calling component.

Now the *termination* and *reachability* objectives for RMCs can be defined. For a vertex $v \in V_i$ and an exit node $ex \in Ex_i$ in the same component A_i, $q^*_{(v,ex)}$ denotes the probability of the global Markov chain reaching state $\langle e, ex \rangle$, starting in state $\langle e, v \rangle$. Then, let $q^*_v = \sum_{ex \in Ex_i} q^*_{(v,ex)}$ be the probability of *termination* for vertex v, i.e., starting at initial state $\langle e, v \rangle$ the probability of the process reaching any exit node in the same component (with empty call stack). The *reachability* probability of vertex v' from vertex v is defined in one of two possible ways:

- either as the probability in the global Markov chain, starting from state $\langle e, v \rangle$, of reaching state $\langle e, v' \rangle$ (i.e., the probability of reaching vertex v' belonging to the same component, with an empty call stack),
- or as the probability in the global Markov chain, starting from state $\langle e, v \rangle$, of reaching state $\langle \beta, v' \rangle$ (i.e., the probability of reaching vertex v' belonging to any component, with some call stack $\beta \in B^*$).

According to [EY09, Proposition 2.1], for any given RMC, either of the two definitions of reachability probability can be expressed in terms of termination probability in a modified RMC, which can be constructed in linear time.

The following are some special subclasses of RMCs: *1-exit* RMCs, where each component has exactly one exit node (but still an arbitrary number of entry nodes)[8]; *1-box* RMCs, where each component contains at most one box inside (1-box RMCs are equivalent to one-counter probabilistic automata); *bounded* RMCs, where the number of components and the number of entry and exit nodes in each of the components are

[8] The restriction of each component having only one entry node is not as interesting, since any multi-entry RMC can be efficiently transformed to an equivalent 1-entry RMC. Hence, one can assume that each component has a single entry. However, the restriction to 1-exit is crucial (e.g., both qualitative and quantitative termination problems have been studied for 1-exit RMCs (and their game generalizations), showing complexity upper bounds, but for multi-exit RMCs even the 1-player generalization is undecidable for these problems [EY15]).

all bounded; *Hierarchical* RMCs, where there can be no cycle of recursive calls among the components.

Recursive Markov Decision Processes (RMDPs), Simple Stochastic Games (RSSGs) and Concurrent Stochastic Games (RCSGs) are the 1-player, 2-player turn-based and 2-player concurrent generalizations of RMCs.

The following two theorems are cited from [EY09], as they bear importance in showing some equivalences between the models defined in this thesis. In particular, the theorems show equivalences between the models with respect to the termination (extinction) objective.

Theorem 2.5 (cf. [EY09], Theorem 2.3)**.**

1. *Every SCFG, $\mathcal{S}$, can be transformed in linear time to a 1-exit RMC, $\mathcal{R}$, such that $|\mathcal{R}| \in O(|\mathcal{S}|)$, and there is a bijection from non-terminals T_j in $\mathcal{S}$ to components A_j of $\mathcal{R}$, each with a single entry en_j and a single exit ex_j, such that $p(T_j) = q^*_{(en_j,ex_j)}$, for all j.*

2. *Conversely, every 1-exit RMC, $\mathcal{R}$, can be transformed in linear time to a SCFG, $\mathcal{S}$, of size $O(|\mathcal{R}|)$, such that there is a map from every vertex u in $\mathcal{R}$ to every non-terminal T_u in $\mathcal{S}$, such that $q^*_u = p(T_u)$.*

Theorem 2.6 (cf. [EY09], Theorem 2.4)**.**

1. *Every BP, $\mathcal{B}$, (even when the BP's rules are presented by giving the multi-sets in a binary representation) can be transformed in polynomial time to a 1-exit RMC, $\mathcal{R}$, such that there is a mapping from types T_j in $\mathcal{B}$ to components A_j of $\mathcal{R}$, each with a single entry en_j and a single exit ex_j, such that the probability, starting at an object of type T_j, of extinction in $\mathcal{B}$ is indeed $= q^*_{(en_j,ex_j)}$, for all j.*

2. *Conversely, every 1-exit RMC, $\mathcal{R}$, can be transformed in linear time to a BP, $\mathcal{B}$, of size $O(|\mathcal{R}|)$, such that there is a map from vertices u in $\mathcal{R}$ to types T_u in $\mathcal{B}$, such that q^*_u is indeed equal to the probability of extinction in $\mathcal{B}$, starting at an object of type T_u.*

Theorems 2.5 and 2.6 above connect the model of RMCs to the models defined in the previous sections in this chapter. In particular, computing the termination probabilities of a 1-exit RMC is equivalent to computing the extinction probabilities of a corresponding BP (see Theorem 2.6) and is also equivalent to computing the probabilities of the generated language (i.e., the termination probabilities) of a corresponding SCFG (see Theorem 2.5). And, as already shown in subsection 2.4.1, computing the extinction probabilities of a BP is equivalent to computing the termination probabilities of a corresponding OBP.

2.6 Related work

This section provides a survey discussion of past work on the models defined in the previous sections, related to the objectives analysed in the thesis. It also discusses similarities and differences between all these models with respect to the properties of (single-target) reachability, extinction/termination and multi-objective reachability.

2.6.1 Single-target reachability objective

As mentioned before, BCSGs is a class of infinite-state imperfect-information stochastic games, that generalize both finite-state concurrent stochastic games and branching simple (turn-based) stochastic games.

The finite-state CSG model was studied in [dAHK07], giving P-time (more precisely, quadratic time) algorithms for the qualitative (single-target) reachability analysis, both for the almost-sure and the limit-sure reachability problems. It was shown that when almost-sure reachability is achieved, the winning player has a randomized memoryless winning strategy, and otherwise the adversary has a "spoiling" randomized strategy that forces reachability probability < 1 against any maximizer strategy (i.e., any strategy for the player maximizing the reachability probability) and that depends only on the number of steps in the game so far. In the case when limit-sure reachability is achieved, the winning player has a family of randomized memoryless winning strategies (one for each $\varepsilon > 0$), and otherwise the adversary has a spoiling randomized memoryless strategy that ensures the reachability probability is upper bounded by some constant. All strategies were shown to be computable in quadratic time in the size of the game and cannot be deterministic in general. In fact, Figure 1 in [dAHK07] shows a simple example that act as an explanation to why deterministic strategies are no longer sufficient for concurrent games. In the concurrent setting, randomization is needed in order to postpone a player's move being revealed (to the other player) until after it is played. Another important observation regarding even finite-state concurrent game settings is that probabilistic states do not contribute to the hardness of the model and can be "simulated" by controlled concurrent states, so the stochastic and non-stochastic variants of the concurrent game model are equivalent, both with respect to computing the game value and to constructing (ε-)optimal strategies for the players (see [EY08, Proposition 2.1]).

Next, BMDPs and BSSGs with (single-target) reachability objective were studied in [ESY18]. It was shown that in a BSSG the player minimizing the reachability probability always has a deterministic static optimal strategy, whereas (unlike for the extinction objective) in general there need not exist *any* optimal strategy for the player maximizing the reachability probability even in a BMDP (and hence also in a BSSG and a BCSG). On the other hand, it was shown in [ESY18] that for BMDPs and BSSGs, if the reachability game value is $= 1$, then there is in fact an optimal strategy (but not in general a static one, even when randomization is allowed) for the player maximizing the reachability probability that forces the value 1 (irrespective of the strategy of the player minimizing the reachability probability). In other words, almost-sure and limit-sure reachability problems coincide for BSSGs (which, as shown later in Chapter 3, is not the case for the more general model of BCSGs). It was also shown that whether the value $= 1$ for BSSG reachability games can be decided in P-time, and if the answer is "yes" then an optimal (non-static, but deterministic) strategy that achieves reachability value 1 for the maximizer can be computed in P-time, whereas if the answer is "no" a deterministic static strategy that forces value < 1 can be computed for the minimizer in P-time.

The study ([ESY18]) also gave polynomial time algorithms for the approximate quantitative reachability analysis of BMDPs, i.e., for computing for a given $\varepsilon > 0$ the $\varepsilon-$optimal reachability probability for maximizing and minimizing BMDPs (together with deterministic static ε-optimal strategies in the case of minimizing BMDPs and randomized static ε-optimal strategies in the case of maximizing BMDPs), and showed that this problem for BSSGs is in TFNP. Note that the problem of exactly computing the reachability value of the BSSG game is at least as hard as the problem of exactly computing the reachability value for finite-state simple stochastic games, where the latter problem is in PLS $\cap$ PPAD and its decision version is the well-known long-standing open problem (called *Condon's* problem) of whether it can be done efficiently in polynomial time [Con92] (it is only known to be in NP $\cap$ coNP)[9]. It was also

[9] It is well-known that the reachability probabilities in finite-state simple stochastic games are obtained as the least fixed point of a system of corresponding Bellman optimality equations ([Con92]).

shown in [ESY18] that the optimal non-reachability probabilities of maximizing or minimizing BMDPs and BSSGs are captured by the greatest fixed point of a system of equations, $x = P(x)$, where the right-hand side $P_i(x)$ of each equation is the maximum or minimum of a set of probabilistic polynomials in x (note that these types of equation systems are special cases of the minimax-PPS, and correspond to the case where in each one-shot matrix game on the right-hand side of the minimax-PPS equations only one of the two players has a choice of actions).

As shown in Proposition 2.4, OBPs are equivalent to BPs with respect to the (single-target) reachability objective (similarly for the MDP and game generalizations of the two models). So all reachability results for BPs (and their 1- and 2-player-controlled extensions) apply, *mutatis mutandis*, to OBPs (and their 1- and 2-player-controlled extensions). It was shown in [BBFK08] that almost-sure single-target reachability in 1-exit RMDPs, or equivalently in Context-Free MDPs with *left-most* derivation, can be decided in polynomial time. However, Context-Free MDPs with *left-most* derivation are very different than (O)BMDPs, which allow *simultaneous* derivation of the tree from all unexpanded non-terminals in each generation (not just the left-most one). Indeed, unlike single-target reachability for OBMDPs (equivalently, Context-Free MDPs with *simultaneous* derivation), even for single-target reachability for 1-exit RMDPs (equivalently, Context-Free MDPs with *left-most* derivation), almost-sure $\neq$ limit-sure (i.e., almost-sure and limit-sure problems do not coincide) and the decidability of limit-sure reachability of a given target remains an open question (despite the fact that there is a polynomial time algorithm for almost-sure reachability).

The quantitative problem for finite-state CSG reachability games, i.e., computing or approximating the value of the game, has been studied previously and seems to be considerably harder than the qualitative problem. The problem of determining whether the value exceeds a given rational number, for example $1/2$, is at least as hard as the long-standing SQRT-SUM problem ([EY08, Theorem 5.1]), mentioned in Section 2.1. The problem of approximating the value within a given desired precision can be solved however in the polynomial hierarchy, specifically in TFNP[NP] ([FM13, Theorem 1]). It is open whether the approximation problem is in NP (or moreover in P). It was shown in [HIJM14] that the standard algorithms for (approximately) solving these games,

The results in [Con92] are formulated in terms of finite-state stopping (i.e., halting with probability 1) SSGs, but the reachability problem for general finite-state SSGs can be in P-time reduced to the reachability problem for finite-state stopping SSGs (showed in [Con92]). In fact, for a stopping SSG the corresponding system of reachability optimality equations has a unique fixed point. Furthermore, such systems of equations are special restricted versions of the systems of equations related to the reachability problem in finite-state concurrent stochastic games (shown how to construct in [Sha53]).

value iteration and *policy iteration*, can take in the worst case a doubly-exponential number of iterations to obtain any nontrivial approximation, even when the reachability value is 1. Note also that there are finite-state CSGs, with reachability value $= 1$, for which (near-)optimal strategies for minimizer (maximizer, respectively) need to have some action probabilities that are doubly-exponentially small (see [CHIJ17] and the finite-state CSG reachability game *Purgatory* in [HKM09]); thus a fixed point representation of the probabilities would need an exponential number of bits, and one must use a suitable compact representation to ensure polynomial space.

2.6.2 Termination (Extinction) objective

Another important objective, namely the probability of termination (extinction), has been studied previously for all these models. The models of BPs and OBPs (and their MDP and game generalizations) under the extinction and termination, respectively, objective are equivalent to corresponding subclasses of Recursive Markov models, called 1-exit Markov Chains (1-RMCs), Markov Decision Processes (1-RMDPs), Simple Stochastic Games (1-RSSGs), and Recursive Concurrent Stochastic Games (1-RCSGs), and related subclasses of probabilistic pushdown processes, under the termination objective [ESY17, ESY20, EY09, EY15, EY08, EKM06]. The models of BPs and OBPs under the extinction (respectively, termination) objective are also equivalent to corresponding models of Stochastic Context-Free Grammars (SCFGs) under the objective of computing the probability of the generated language (see subsection 2.5.1 for more details). More precisely, there are pairwise reductions between the following problems and, hence, they are equivalent: (1) computing the termination probabilities of a 1-exit RMC; (2) computing the extinction probabilities of a BP; (3) computing the probabilities of the generated language of a SCFG; and (4) computing the termination probabilities of a OBP (similarly, for the MDP and game generalizations of all these models). Theorems 2.5 and 2.6 (cited from study [EY09]) showed the equivalences (1) $\leftrightarrow$ (2), (1) $\leftrightarrow$ (3) and (2) $\leftrightarrow$ (3) with respect to the purely probabilistic setting, but it is not difficult to generalize the theorems to the MDP and game variants. And although it is easy to see, for completeness subsection 2.4.1 gave the equivalence (2) $\leftrightarrow$ (4).

The extinction (or termination) probabilities for all these models are captured by the *least fixed point* (LFP) solutions of similar systems of probabilistic polynomial equations. For example, the extinction values of a BCSG (equivalently, the termination values of an 1-RCSG, an OBCSG and a CF-CSG) are given by the LFP of a minimax-

PPS (though not the exactly the same minimax-PPS as for the (non)-reachability objective). This was shown in [EY08].

Polynomial time-algorithms for the qualitative analysis, as well as for the approximate computation of the extinction probabilities for BPs (equivalently, termination probabilities for 1-RMCs, SCFGs and OBPs) were given in [EY09, ESY17, EGK10]. For optimal extinction probabilities in BMDPs (equivalently, optimal termination probabilities in 1-RMDPs, CF-MDPs and OBMDPs), P-time algorithms for the qualitative decision problems and for the quantitative approximation problems (both for a maximizing and minimizing player), for both computing the ε-optimal probabilities and an ε-optimal static strategy, were shown in [EY06, EY15, ESY20]. However, negative results were shown which indicate that the problem is much harder for branching concurrent (or even simple) stochastic games, even for the qualitative extinction problem. Specifically, it was shown in [EY06, EY15] that the qualitative extinction problem for BSSGs (equivalently, the qualitative termination problem for 1-RSSGs, CF-SSGs and OBSSGs) is in NP$\cap$coNP and is at least as hard as the well-known *Condon's* quantitative open problem for the value of a finite-state simple stochastic game ([Con92]). Also, [EY15] showed that both the maximizer and minimizer of the extinction (termination) probability always (not only when the value is 1) have an optimal static deterministic strategy. Furthermore, it was shown in [EY08] that (both the almost-sure and limit-sure) qualitative extinction problems for BCSGs (equivalently, qualitative termination problems for 1-RCSGs, CF-CSGs and OBCSGs) are at least as hard as the SQRT-SUM problem (which is not known to be even in the Polynomial Hierarchy, for more details on the SQRT-SUM problem refer to Section 2.1).[10] It was also shown in [EY08], using a strategy iteration method, that the player minimizing the extinction (termination) probability always has an optimal randomized static strategy, whereas the player maximizing the extinction (termination) probability in general may only have ε-optimal randomized static strategies, for all $\varepsilon > 0$.

For the quantitative extinction problems for BPs (equivalently, the quantitative termination problems for 1-RMCs, SCFGs and OBPs) and their MDPs and game gener-

[10] The results in [EY08] were phrased in terms of the limit-sure problem, where it was shown that (a) deciding whether the value of a finite-state CSG reachability game is at least a given value $p \in (0,1)$ is SQRT-SUM-hard, and (b) that the former problem is reducible to the limit-sure decision problem for BCSG extinction games. But the hardness proofs of (b) and (a) in [EY08] apply *mutatis mutandis* to (b) the almost-sure decision problem for BCSG extinction, and to (a) the corresponding problem of deciding, given a finite-state CSG and a value $p \in (0,1)$, whether the maximizing player has a strategy that achieves at least value p, regardless of the strategy of the minimizer. Thus, both the almost-sure and limit-sure extinction problems for BCSGs are SQRT-SUM-hard, and also both are at least as hard as *Condon's* problem of computing the exact value of a finite-state SSG reachability game.

alizations, all decision problems (and thus, all approximation problems where a P-time algorithm has not been shown yet) have PSPACE as upper bound, relying on the upper bounds for the decision procedures for ETR ($\exists \mathbb{R}$). Moreover, such decision problems are PosSLP-hard (and thus, SQRT-SUM-hard) ([EY09, Theorem 5.1, 5.3]).

The equivalence between the models of BPs and OBPs (and their MDP and game generalizations) and the model of 1-RMCs (and their MDP and game generalizations) with respect to extinction (termination) does not hold for the (single-target) reachability objective. For example, almost-sure and limit-sure reachability problems coincide for (O)BMDPs, i.e., if the supremum probability of reaching the target is 1 then there exists a strategy that ensures reachability with probability exactly 1. However, this is not the case for 1-RMDPs. Furthermore, as mentioned in 2.6.1, it is known that almost-sure reachability for 1-RMDPs can be decided in polynomial time ([BBFK08, BBKO11]), but limit-sure reachability for 1-RMDPs is not even known to be decidable. The qualitative reachability problem for 1-RMDPs and 1-RSSGs (and equivalent probabilistic pushdown models) was studied in [BBKO11, BKL14]. These results do not apply to the corresponding models of (O)BMDPs and (O)BSSGs. For an extensive survey on questions, such as extinction/termination and reachability, on probabilistic pushdown automata (which under specific restrictions can be equivalent to SCFGs, BPs and RMCs), please see [BEKK13].

2.6.3 Multi-objective reachability

Multi-objective reachability and model checking (with respect to omega-regular properties) have been studied for finite-state MDPs in [EKVY08], both with respect to qualitative and quantitative problems. In particular, it was shown in [EKVY08] that for multi-objective reachability in finite-state MDPs, memoryless (but randomized) strategies are sufficient, that both qualitative and quantitative multi-objective reachability queries can be decided in P-time, and that the *Pareto curve* for them can be approximated within a desired error $\varepsilon > 0$ in P-time in the size of the MDP and $1/\varepsilon$.

As pointed out in the next subsection, model checking for the infinite-state model of Branching Processes has been studied before. However, to the best of our knowledge, there are no previous results on multi-objective reachability for BMDPs (and also OBMDPs, which is a model that has been only recently defined in [EM20], which is one of the core papers that this thesis is written on), neither for the qualitative nor for the quantitative multi-objective reachability questions that are defined in Section 2.4.

2.6.4 Richer objectives

Another objective considered in prior work is the *expected total reward* objective for 1-RSSGs ([EWY19]) and 1-RCSGs ([Woj13]) with positive rewards. In particular, [Woj13] shows that the “qualitative” problem of determining whether the game value for a 1-RCSG total reward game is $= \infty$ is in PSPACE. None of these prior results have any implications for BCSGs with reachability objectives.

For richer objectives beyond reachability or extinction, [CDK12] studied model checking of purely stochastic BPs with respect to properties expressed by a deterministic parity tree automaton (DPTA). That is, [CDK12] studied problems of, given a BP and a DPTA, computing the probability that the generated tree in the BP is accepted by the DPTA. It showed that the qualitative problem is in P-time (hence this holds in particular for reachability probability in BPs), and that the quantitative problem of comparing the probability with a rational is in PSPACE (by similarly expressing the problem in terms of a system of non-linear probabilistic polynomial equations and relying on the upper bounds for decision procedures of ETR ($\exists \mathbb{R}$)). Then [MM15] extended this to properties of BPs expressed by “game automata”, a subclass of alternating parity tree automata. More recently, [PS16] considered determinacy (i.e., existence of game values) and complexity of decision problems for ordered branching simple (i.e., turn-based) stochastic games, with regular objectives, where the two players aim to maximize/minimize the probability that the generated labeled tree belongs to a regular language (given by a finite tree automaton). They showed that (unlike the case of games with a simpler objective like reachability) already for some basic regular properties these games are not even determined, meaning they do not have a value. They furthermore showed that for what amounts to OBMDPs with a regular tree objective it is undecidable to compare the optimal probability to a threshold value; whereas for deterministic turn-based branching games they showed it is decidable and 2-EXPTIME-complete (respectively, EXPTIME-complete), to determine whether the player aiming to satisfy (respectively, falsify) a given regular tree objective has a pure winning strategy. Other past research includes work in operations research on Branching MDPs (see e.g. [Pli76, RW82, DR05]). None of these prior works on richer objectives bear on any of the results established in this thesis.

Chapter 3

Branching Concurrent Stochastic Games

In this chapter we focus on the BCSG game model with respect to the *(single-target) reachability objectives*, a basic and natural class of objectives. Some types are designated as undesirable (for example, malignant cells in cancer), in which case we want to minimize the probability of ever reaching any object of such a type. Or conversely, some types may be designated as desirable, in which case we want to maximize the probability of reaching an object of such a type.

First, a summary of the main results of this chapter. We first show that a BCSG with a reachability objective has a well-defined value, i.e., given an initial (finite) population μ of objects of various types and a target type T_{f^*}, if the sets of (mixed) strategies of the two players are respectively Ψ_1, Ψ_2, and if $\Upsilon_{\sigma,\tau}(\mu, T_{f^*})$ denotes the probability of reaching eventually an object of type T_{f^*} when starting from population μ under strategy $\sigma \in \Psi_1$ for player 1 and strategy $\tau \in \Psi_2$ for player 2, then $\inf_{\sigma\in\Psi_1} \sup_{\tau\in\Psi_2} \Upsilon_{\sigma,\tau}(\mu, T_{f^*}) = \sup_{\tau\in\Psi_2} \inf_{\sigma\in\Psi_1} \Upsilon_{\sigma,\tau}(\mu, T_{f^*})$, which is the value v^* of the game. Furthermore, we show that the player who wants to minimize the reachability probability always has an optimal (mixed) static strategy that achieves the value, i.e., a strategy σ^* which uses for all objects of each type T generated over the whole history of the game the same probability distribution on the available actions, independent of the past history, and which has the property that $v^* = \sup_{\tau\in\Psi_2} \Upsilon_{\sigma^*,\tau}(\mu, T_{f^*})$. The optimal strategy in general has to be mixed (randomized), since this was known to be the case even for finite-state concurrent games (see [dAHK07, Figure 1]). On the other hand, the player that wants to maximize the reachability probability of a BCSG may not have an optimal strategy (whether static or not), and it was known that this holds even for BMDPs, i.e., even when there is only one player (see [ESY18, Example 3.2]). This also holds for finite-state CSGs: the player aiming to maximize the reachability probability does not necessarily have any optimal strategy (shown in [dAHK07, Figure 2] to be true even when the reachability value is 1).

As mentioned in the previous chapter, to analyze BCSGs with respect to the reachability objective, we model them by a system of equations $x = P(x)$, called a *minimax Probabilistic Polynomial System* (minimax-PPS for short), where x is a tuple of variables corresponding to the types of the BCSG. There is one equation, $x_i = P_i(x)$, for each type T_i, where $P_i(x)$ is the minimax value of a (one-shot) two-player zero-sum matrix game, whose payoff for every pair of actions is given by a polynomial in x whose coefficients are positive and sum to at most 1 (a probabilistic polynomial). The function $P(x)$ defines a monotone operator from $[0,1]^n$ to itself, and thus it has, in particular, a *greatest fixed point* (GFP) g^* in $[0,1]^n$. We show that the coordinates g^*_i of the GFP give the *non-reachability* values for the BCSG game when started with a population that consists of a single object of type T_i.[1] The value of the game for any initial population μ can be derived easily from the GFP, g^*, of the minimax-PPS. This generalizes the result in [ESY18, Theorem 3.1], which established an analogous result for the special case of BSSGs. It also follows from our minimax-PPS equational characterization that *quantitative* decision problems for BCSGs, such as deciding whether the reachability game value is $\geq p$ for a given $p \in (0,1)$ are all solvable in PSPACE.

Our main algorithmic results in this chapter concern the qualitative analysis of the reachability problem, that is, the problem of determining whether one of the players can win the game with probability 1, i.e., if the value of the game is 0 or 1. We provide the first polynomial time algorithms for qualitative reachability analysis for BCSGs. For the value $= 0$ problem, the algorithm and its analysis are rather simple. If the value is 0, the algorithm computes an optimal strategy σ^* for the player that wants to minimize the reachability probability; the constructed strategy σ^* is in fact static and deterministic, i.e., it selects for each type deterministically a single available action, and guarantees $\Upsilon_{\sigma^*,\tau}(\mu, T_{f^*}) = 0$ for all $\tau \in \Psi_2$. If the value is positive then the algorithm computes a static randomized strategy τ for the player maximizing the reachability probability that guarantees $\inf_{\sigma\in\Psi_1} \Upsilon_{\sigma,\tau}(\mu, T_{f^*}) > 0$.

[1] As a comparison for a more complete picture, if there were no variables (types) of Q-form, the equations would be precisely the Bellman optimality equations for finite-state CSGs with the objective of non-reachability. It is a well-known fact that for finite-state CSGs the reachability values are obtained from the *least fixed point* of a system of Bellman optimality equations. In other words, the equations in the minimax-PPSs can be seen as a generalized version of Shapley's optimality equations for finite-state CSGs.

The value $= 1$ problem is much more complicated. There are two versions of the value $= 1$ problem, because it is possible that the value of the game is 1 but there is no strategy for the maximizing player that guarantees reachability with probability exactly 1 (see Example 3.1). The critical reason for this is the concurrency in the moves of the two players: for BMDPs and BSSGs, it is known that if the value is 1 then there is a strategy τ that achieves it ([ESY18, Theorem 9.4])[2]; on the other hand, this is not the case even for finite-state CSGs ([dAHK07, Figure 2]). Thus, we have two versions of the problem. In the first version, called the *almost-sure problem*, we want to determine whether there exists a strategy τ^* for player 2 that guarantees that the target type T_{f^*} is reached with probability 1 regardless of the strategy of player 1, i.e., such that $\Upsilon_{\sigma,\tau^*}(\mu, T_{f^*}) = 1$ for all $\sigma \in \Psi_1$. In the second version of the problem, called the *limit-sure problem*, we want to determine if the value $v^* = \sup_{\tau\in\Psi_2} \inf_{\sigma\in\Psi_1} \Upsilon_{\sigma,\tau}(\mu, T_{f^*})$ is 1, i.e., if for every $\varepsilon > 0$ there is a strategy τ_ε for player 2 that guarantees that the probability of reaching the target type is at least $1-\varepsilon$ regardless of the strategy σ for player 1; such a strategy τ_ε is called $\varepsilon - optimal$. The main results of the chapter are to provide polynomial time algorithms for both versions of the problem. The algorithms are nontrivial, building upon the algorithms of both [dAHK07] and [ESY18] which both address different special subcases of qualitative BCSG reachability.

In the almost-sure problem, if the answer is positive, our algorithm constructs (a compact description of) a strategy τ^* for player 2 that achieves value 1; the strategy is a randomized non-static strategy, and this is inherent (i.e., there may not exist a static strategy that achieves value 1). If the answer is negative, then our algorithm constructs a randomized non-static strategy σ for the opposing player 1 such that $\Upsilon_{\sigma,\tau}(\mu, T_{f^*}) < 1$ for all strategies τ of player 2. In the limit-sure problem, if the answer is positive, i.e., the value is 1, our algorithm constructs for any given $\varepsilon > 0$, a randomized static ε-optimal strategy, i.e., a strategy $\tau_\varepsilon \in \Psi_2$ such that $\Upsilon_{\sigma,\tau_\varepsilon}(\mu, T_{f^*}) \geq 1-\varepsilon$ for all $\sigma \in \Psi_1$. If the answer is negative, i.e., the value is < 1, our algorithm constructs a randomized static strategy σ' for player 1 such that $\sup_{\tau\in\Psi_2} \Upsilon_{\sigma',\tau}(\mu, T_{f^*}) < 1$.

Finally, we discuss the complexity of BPs (and their MDP and game generalizations) with respect to the reachability objective. By adapting analogous results from previous papers on the model of Recursive Markov chains (namely, [EY08, Theorem 3.3] and [EY09, Theorem 5.3]), we provide for completeness the PSPACE upper bound for both quantitative reachability decision and approximation questions, and the

[2]When the value is positive and *not* equal to 1, even for BMDPs there need not exist an optimal strategy for the player maximizing the reachability probability (see [ESY18, Example 3.2]).

PosSLP lower bound for the decision questions. We also show that computing the optimal reachability probabilities for minimizing BMDPs, equivalently computing the GFP in a maxPPS, is in FIXP.

Organization of the chapter. Section 3.1 shows the relationship between the non-reachability values of a BCSG game and the greatest fixed point of a minimax-PPS. Section 3.2 presents the algorithm for determining if the reachability value of a BCSG is 0. Section 3.3 shows some preliminary results for (minimax-)PPSs needed for the discussion of the value $= 1$ case. Section 3.4 presents the algorithm for almost-sure reachability, and Section 3.5 for limit-sure reachability. Finally, Section 3.6 finishes with upper and lower bounds discussion for reachability in BPs and their MDPs and (concurrent) game variants.

3.1 Non-reachability values for BCSGs and the Greatest Fixed Point

This section will show that for a given BCSG with a target type T_{f^*}, a minimax-PPS, $x = P(x)$, can be constructed such that its *Greatest Fixed Point* (GFP), $g^* \in [0,1]^n$, is precisely the vector g^* of non-reachability values for the BCSG.

For simplicity, from now on let us call a *maximizer* (respectively, a *minimizer*) the player that aims to *maximize* (respectively, *minimize*) the probability of *not* reaching the target type. That is, we swap the roles of the players for the benefit of less confusion in analysing the minimax-PPS. While the players' goals in the game are related to the objective of reachability, the equations we construct will capture the optimal non-reachability values in the GFP of the minimax-PPS.

For each type $T_i \neq T_{f^*}$, the minimax-PPS will have an associated variable x_i and an equation $x_i = P_i(x)$, and the MINIMAX-PROBABILISTIC-POLYNOMIAL $P_i(x)$ is built in the following way. For each action $a_{max} \in \Gamma^i_{max}$ of the maximizer (i.e., the player aiming to maximize the probability of *not* reaching the target) and action $a_{min} \in \Gamma^i_{min}$ of the minimizer in T_i, let $R'(T_i, a_{max}, a_{min}) = \{r \in R(T_i, a_{max}, a_{min}) \mid (\alpha_r)_{f^*} = 0\}$ be the set of probabilistic rules r for type T_i and players' action pair (a_{max}, a_{min}) that generate a multi-set α_r which does not contain an object of the target type. For each actions pair for T_i, there is a probabilistic polynomial $q_{i,a_{max},a_{min}}(x) := \sum_{r \in R'(T_i,a_{max},a_{min})} p_r x^{\alpha_r}$. Observe that there is no need to include rules where α_r contains an object of type

T_{f^*}, because then the term with monomial x^{α_r} will be 0. Now after a polynomial is constructed for each pair of players' moves, we construct $P_i(x)$ as the minimax value of a zero-sum matrix game $A_i(x)$ (i.e., $P_i(x) := Val(A_i(x))$), where the matrix is constructed as follows: (1) rows belong to the max player in the minimax-PPS (i.e., the player trying to maximize the non-reachability probability), and columns belong to the min player; (2) for each row and column (i.e., pair of actions (a_{max}, a_{min})) there is a corresponding probabilistic polynomial $q_{i,a_{max},a_{min}}(x)$ in the matrix entry $A_i(x)_{(a_{max},a_{min})}$.

The following theorem captures the fact that the optimal *non-reachability values* g^* in the BCSG correspond to the *Greatest Fixed Point (GFP)* of the minimax-PPS.

Theorem 3.1. *The non-reachability game values $g^* \in [0,1]^n$ of a BCSG reachability game exist, and correspond to the Greatest Fixed Point (GFP) of the minimax-PPS, $x = P(x)$, in $[0,1]^n$. That is, $g^* = P(g^*)$, and for all other fixed points $g' = P(g')$ in $[0,1]^n$, it holds that $g' \leq g^*$. Moreover, for an initial population μ, the optimal non-reachability value is $g^*(\mu) = \prod_i (g_i^*)^{(\mu)_i}$ and the game is determined, i.e., $g^*(\mu) = \sup_{\sigma\in\Psi_1} \inf_{\tau\in\Psi_2} g^*_{\sigma,\tau}(\mu) = \inf_{\tau\in\Psi_2} \sup_{\sigma\in\Psi_1} g^*_{\sigma,\tau}(\mu)$. Finally, the player maximizing non-reachability probability in the BCSG has a (mixed) static optimal strategy.*

Proof. Note that $P : [0,1]^n \to [0,1]^n$ is a monotone operator, since all coefficients in all the polynomials $P_i(x)$ are non-negative, and for $x \leq y$, where $x, y \in [0,1]^n$, it holds that $A_i(x) \leq A_i(y)$ (entry-wise inequality) and thus $Val(A_i(x)) \leq Val(A_i(y))$, where recall $Val()$ is the minimax value operator. Thus, $P_i(x) \leq P_i(y)$. Let $x^0 = \mathbf{1}$ and $x^k = P(x^{k-1}) = P^k(\mathbf{1})$, $k > 0$ be the k-fold application of P on the vector $\mathbf{1}$ (i.e., the all-1 vector). By induction on k the sequence x^k is monotonically non-increasing, i.e., $x^{k+1} \leq x^k \leq \mathbf{1}$ for all $k > 0$.

By Tarski's theorem ([Tar55, Theorem 1]), $P(\cdot)$ has a Greatest Fixed Point (GFP) $x^* \in [0,1]^n$. The GFP is the limit of the monotone the sequence x^k, i.e., $x^* = \lim_{k\to\infty} x^k$. To continue the proof, we need the following lemma.

Lemma 3.2. *For any initial non-empty population μ, assuming it does not contain the target type T_{f^*}, and for any $k \geq 0$, the value of not reaching T_{f^*} in k steps is $g^k(\mu) = f(x^k, \mu) := \prod_{i=1}^n (x_i^k)^{(\mu)_i}$. Also, there are strategies for the players, $\sigma^k \in \Psi_1$ and $\tau^k \in \Psi_2$, that achieve this value, that is $g^k(\mu) = \sup_{\sigma\in\Psi_1} g^k_{\sigma,\tau^k}(\mu) = \inf_{\tau\in\Psi_2} g^k_{\sigma^k,\tau}(\mu)$.*

Proof. Before we begin the proof, let us make a quick observation. For a fixed vector $x \in [0,1]^n$, consider the zero-sum matrix game defined by the payoff matrix $A_i(x)$ for player 1 (the row player). Consider fixed mixed strategies $\mathbf{s}_i$ and $\mathbf{t}_i$ for the row and column players in this matrix game. Thus, $\mathbf{s}_i(a_{max})$ ($\mathbf{t}_i(a_{min})$, respectively) defines the probability placed on action $a_{max} \in \Gamma^i_{max}$ (on action $a_{min} \in \Gamma^i_{min}$, respectively) in $\mathbf{s}_i$ (in $\mathbf{t}_i$, respectively). The expected payoff to player 1 (the maximizing player), under these mixed strategies is:

$$\begin{aligned}
&\sum_{a_{max}\in\Gamma^i_{max},a_{min}\in\Gamma^i_{min}} \mathbf{s}_i(a_{max})\mathbf{t}_i(a_{min})q_{i,a_{max},a_{min}}(x) \\
=& \sum_{a_{max},a_{min}} \left[\mathbf{s}_i(a_{max})\mathbf{t}_i(a_{min}) \sum_{r\in R'(T_i,a_{max},a_{min})} p_r x^{\alpha_r}\right] \\
=& \sum_{a_{max},a_{min}} \sum_{r\in R'(T_i,a_{max},a_{min})} \mathbf{s}_i(a_{max})\mathbf{t}_i(a_{min})p_r x^{\alpha_r} = \sum_{r\in R'(T_i)} p'_r x^{\alpha_r} \qquad (3.1)
\end{aligned}$$

where $R'(T_i)$ is the set of all probabilistic rules for type T_i; the newly defined probability p'_r of a rule r is equal to $\mathbf{s}_i(a_{max}) * \mathbf{t}_i(a_{min}) * p_r$ for the pair (a_{max}, a_{min}) for which the rule r is in $R'(T_i, a_{max}, a_{min})$, and where α_r is the population that rule r generates, meaning rule r is defined by $T_i \xrightarrow{p_r} \alpha_r$.

Now let us prove the Lemma by induction on k. For the basis step, clearly $g^0(\mu) = \mathbf{1}$, since the initial population does not contain any objects of the target type. Moreover, $x^0 = \mathbf{1}$ and so $f(\mathbf{1},\mu) = \mathbf{1}$.

For the inductive step, first we demonstrate that $g^k(\mu) \geq f(x^k,\mu)$. Consider a strategy $\sigma^k := (\hat{\mathbf{s}}, \sigma^{k-1})$ for the max player (i.e., the player aiming to maximize the non-reachability probability), constructed in the following way. For all i, and for every object of type T_i in the initial population $\mu = X_0$, the max player chooses as a first step the minimax-optimal mixed strategy $\hat{\mathbf{s}}_i$ in the zero-sum matrix game $A_i(x^{k-1})$ (which exists, due to the minimax theorem). The min player (player 2), as part of its strategy, chooses some distributions on actions for all objects in the population X_0 (independently of player 1), and then the rules are chosen according to the resulting probabilities, forming the next generation X_1 at time 1. Thereafter, the max player acts according to an optimal $(k-1)$-step strategy σ^{k-1}, starting from population X_1 (σ^{k-1} exists by the inductive assumption, and we will indeed prove by induction that the thus defined k-step strategy σ^k is optimal in the k-step game). Note that σ^k can be mixed, and can also be non-static since the action probabilities can depend on the generation and history.

Now let τ be any strategy for the min player. In the first step, τ chooses some distributions on actions for each object in $X_0 = \mu$. After the choices of σ^k and τ are made in the first step, rules are picked probabilistically and the population X_1 is generated. By the inductive assumption, $g^{k-1}(X_1) = f(x^{k-1},X_1)$, i.e., the value of not reaching the target type in next $k-1$ steps, starting in population X_1, is precisely $f(x^{k-1}, X_1)$. Therefore, the k-step probability of not reaching the target, starting in μ, using strategies σ^k and τ, is $g^k_{\sigma^k,\tau}(\mu) = \sum_{X_1} p(X_1) g^{k-1}_{\sigma^{k-1},\tau}(X_1) \geq \sum_{X_1} p(X_1) f(x^{k-1}, X_1)$, where the sum is over all possible next-step populations X_1, and in each term $f(x^{k-1}, X_1)$ is multiplied by the probability $p(X_1)$ of generating that particular population X_1. The reason for the inequality is because, by optimality of σ^{k-1} for the max player in the $(k-1)$-step game, we know that $g^{k-1}_{\sigma^{k-1},\tau}(X_1) \geq \inf_{\pi \in \Psi_2} g^{k-1}_{\sigma^{k-1},\pi}(X_1) = g^{k-1}_{\sigma^{k-1},*}(X_1) = g^{k-1}(X_1) = f(x^{k-1}, X_1)$.

The sum $\sum_{X_1} p(X_1) f(x^{k-1}, X_1)$ can be rewritten as a product of $|\mu|$ terms, one for each object in the initial population X_0, where for a n-vector $\mu \in \mathbb{N}^n$, let $|\mu|$ denote the L^1-norm of vector μ, i.e., $|\mu| := \sum_{i=1}^n (\mu)_i$. Specifically, given X_0, let L_{X_0,X_1} denote the set of all possible tuples of rules $(r_1, \ldots, r_{|X_0|})$, which associate to each object e_j in the population X_0, a rule r_j such that if e_j has type T_i, then $r_j \in R'(T_i)$ is a rule for type T_i, and furthermore such that if we apply the rules $(r_1, \ldots, r_{|X_0|})$, they generate multi-sets $\alpha_1, \ldots, \alpha_{|X_0|}$, such that we obtain the population $X_1 = \bigcup \alpha_i$ from them.

Then for $X_0 = \mu$, we can rewrite $\sum_{X_1} p(X_1) f(x^{k-1}, X_1)$ as:

$$\begin{aligned} \sum_{X_1} p(X_1) f(x^{k-1}, X_1) &= \sum_{X_1} \sum_{(r_1,\ldots,r_{|X_0|}) \in L_{X_0,X_1}} \Big(\prod_{j=1}^{|\mu|} p'_{r_j}\Big) \cdot \Big(\prod_{j=1}^{|\mu|} f(x^{k-1}, \alpha_{r_j})\Big) \\ &= \prod_{j=1}^{|\mu|} \sum_{r_j} p'_{r_j} f(x^{k-1}, \alpha_{r_j}) \end{aligned}$$

where r_j ranges over all rules that can be generated by the type of object e_j, and p'_{r_j} is the probability of generating rule r_j for object e_j in the first step, under strategies σ^k and τ. α_{r_j} is the population produced from e_j under rule r_j. Note that the term $\sum_{r_j} p'_{r_j} f(x^{k-1}, \alpha_{r_j})$ for an object e_j of type T_i has the same form as equation (3.1) above. This observation implies that, since the mixed strategy $\hat{\mathbf{s}}_i$ is minimax-optimal in the zero-sum matrix game with matrix $A_i(x^{k-1})$, the term $\sum_{r_j} p'_{r_j} f(x^{k-1}, \alpha_{r_j})$ corresponding to each object e_j of type T_i is $\geq Val(A_i(x^{k-1})) = P_i(x^{k-1}) = x_i^k$. Hence, for any strategy τ chosen the min player, starting with the objects in $\mu = X_0$, the probability of not reaching the target type in next k steps under strategies σ^k and τ is $g^k_{\sigma^k,\tau}(\mu) \geq \prod_{i=1}^{|\mu|} x_i^k = f(x^k, \mu)$. Therefore, the k-step non-reachability value is $g^k(\mu) = \sup_{\sigma \in \Psi_1} \inf_{\tau \in \Psi_2} g^k_{\sigma,\tau}(\mu) \geq \inf_{\tau \in \Psi_2} g^k_{\sigma^k,\tau}(\mu) \geq f(x^k, \mu)$.

Symmetrically we can prove the reverse inequality by using the other player as an argument. That is, similarly let τ^k select as a first step for each object of type T_i in the initial population $\mu = X_0$ the (mixed) optimal strategy in the corresponding zero-sum matrix game $A_i(x^{k-1})$ (exists by the minimax theorem). Simultaneously and independently the max player chooses moves for the objects, and then rules are picked in order to generate population X_1. Afterwards, the min player acts according to an optimal $k-1$-step strategy τ^{k-1} (which exists by the inductive hypothesis). As before, $g^k(\mu)$ can be written as a product of $|\mu|$ terms, where each term is $\sum_{r_j} p'_{r_j} f(x^{k-1}, \alpha_{r_j})$. Again, by the choice of τ^k, it follows that the term for each object e_j of type T_i is at most $Val(A_i(x^{k-1})) = P_i(x^{k-1}) = x_i^k$. Thus, showing that $\sup_{\sigma \in \Psi_1} g^k_{\sigma,\tau^k}(\mu) \leq f(x^k, \mu)$, and $g^k(\mu) \leq f(x^k,\mu)$. So, at the end $g^k(\mu) = \sup_{\sigma \in \Psi_1} g^k_{\sigma,\tau^k}(\mu) = \inf_{\tau \in \Psi_2} g^k_{\sigma^k,\tau}(\mu) = f(x^k,\mu) = \prod_{i=1}^n (x_i^k)^{(\mu)_i}$. Note that the constructed strategy σ^k (and τ^k) is thus optimal for the player maximizing (respectively, minimizing), the probability of not reaching the target type in k steps. If the initial population consists of a single object of type $T_i \neq T_{f^*}$, then the Lemma states that $g_i^k = x_i^k$ for all $k \geq 0$. □

Now we continue the proof of Theorem 3.1. We show that the game is determined, i.e., $g^*(\mu) = \sup_{\sigma \in \Psi_1} \inf_{\tau \in \Psi_2} g^*_{\sigma,\tau}(\mu) = \inf_{\tau \in \Psi_2} \sup_{\sigma \in \Psi_1} g^*_{\sigma,\tau}(\mu)$, and that the game value for the objective of not reaching T_{f^*} is precisely $f(x^*,\mu)$, where $x^* = \lim_{k \to \infty} x^k \in [0,1]^n$ is the GFP of the system $x = P(x)$, which exists by Tarski's theorem. As a special case, if the initial population μ is just a single object of type $T_i \neq T_{f^*}$, we have $g_i^* = x_i^*$.

Since the sequence x^k converges to x^* monotonically from above (recall $x^0 = \mathbf{1}$ and the sequence is monotonically non-increasing), then $f(x^k,\mu)$ converges to $f(x^*,\mu)$ from above, i.e., for any $\varepsilon > 0$ there is a $k(\varepsilon)$ where $f(x^*,\mu) \leq f(x^{k(\varepsilon)},\mu) < f(x^*,\mu) + \varepsilon$. By Lemma 3.2, the min player strategy $\tau^{k(\varepsilon)}$ (as described in the Lemma) achieves the $k(\varepsilon)$-step value of the game, i.e., $\sup_{\sigma \in \Psi_1} g^{k(\varepsilon)}_{\sigma,\tau^{k(\varepsilon)}}(\mu) = f(x^{k(\varepsilon)},\mu) < f(x^*,\mu) + \varepsilon$. But for any strategy σ, $g^*_{\sigma,\tau^{k(\varepsilon)}}(\mu) \leq g^{k(\varepsilon)}_{\sigma,\tau^{k(\varepsilon)}}(\mu)$, since the more steps the game takes, the lower the probability of non-reachability is. So it follows that $\sup_{\sigma \in \Psi_1} g^*_{\sigma,\tau^{k(\varepsilon)}}(\mu) \leq \sup_{\sigma \in \Psi_1} g^{k(\varepsilon)}_{\sigma,\tau^{k(\varepsilon)}}(\mu) < f(x^*,\mu) + \varepsilon$. And since it holds for every $\varepsilon > 0$, then $\inf_{\tau \in \Psi_2} \sup_{\sigma \in \Psi_1} g^*_{\sigma,\tau}(\mu) \leq f(x^*,\mu)$. Thus, by standard facts, $g^*(\mu) = \sup_{\sigma \in \Psi_1} \inf_{\tau \in \Psi_2} g^*_{\sigma,\tau}(\mu) \leq \inf_{\tau \in \Psi_2} \sup_{\sigma \in \Psi_1} g^*_{\sigma,\tau}(\mu) \leq f(x^*,\mu)$.

To show the reverse inequality, namely $g^*(\mu) \geq f(x^*,\mu)$, let σ^* be the (mixed) static strategy for the max player (i.e., the player aiming to maximize the probability of *not* reaching the target type), that for each object of type T_i always selects the (mixed) optimal strategy in the zero-sum matrix game $A_i(x^*)$ (which exists by the minimax theorem). Fixing σ^*, the BCSG becomes a minimizing BMDP and the minimax-PPS, $x = P(x)$, becomes a minPPS, $x = P'(x) = P_{\sigma^*,*}(x)$. In this new system of equations, for every type T_i (i.e., variable x_i), the function on the right-hand side changes from $P_i(x) =$

$Val(A_i(x))$ to $P'_i(x) = \min\{m_b : b \in \Gamma^i_{min}\}$, where $m_b := \sum_{j \in \Gamma^i_{max}} \sigma^*(x_i, j) * q_{i,j,b}(x)$. Hence, $P'(x) \leq P(x)$ for all $x \in [0,1]^n$. Thus, if we denote by $y^k, k \geq 0$ the vectors obtained from the k-fold application of $P'(x)$ on the vector $\mathbf{1}$ (i.e., the all-1 vector), then $y^k \leq x^k$ for all $k \geq 0$. So it follows that $y^* \leq x^*$, with y^* and x^* being the GFP of $x = P'(x)$ and $x = P(x)$, respectively. But since the fixed strategy σ^* is the optimal strategy for the max player with respect to vector x^* and achieves the value $P_i(x^*) = Val(A_i(x^*))$ for all variables, x^* must also be a fixed point of $x = P'(x)$ and hence $x^* = y^*$.

Now consider any strategy τ for the min player in the minimizing BMDP. Recall that a minimizing BMDP is a BCSG where in every type the max player has a single available action. Then by the induction step in the proof of Lemma 3.2 it holds that for every $k \geq 0$, starting in the initial population μ, the probability of *not* reaching the target type T_{f^*} in k steps under strategy τ is at least $f(y^k, \mu)$. Hence, the infimum probability of *not* reaching the target type (in any number of steps) is at least $\lim_{k \to \infty} f(y^k, \mu) = f(y^*, \mu) = f(x^*, \mu)$. Therefore, $\inf_{\tau \in \Psi_2} g^*_{\sigma^*,\tau}(\mu) \geq f(x^*, \mu)$. However, we know that $g^*(\mu) = \sup_{\sigma \in \Psi_1} \inf_{\tau \in \Psi_2} g^*_{\sigma,\tau}(\mu) \geq \inf_{\tau \in \Psi_2} g^*_{\sigma^*,\tau}(\mu)$, which shows the reverse inequality.

We can deduce that $g^*(\mu) = \sup_{\sigma \in \Psi_1} \inf_{\tau \in \Psi_2} g^*_{\sigma,\tau}(\mu) = \inf_{\tau \in \Psi_2} \sup_{\sigma \in \Psi_1} g^*_{\sigma,\tau}(\mu) = f(x^*, \mu) = \inf_{\tau \in \Psi_2} g^*_{\sigma^*,\tau}(\mu)$ and σ^* is an optimal (mixed) static strategy for the max player under the non-reachability objective. □

Note that the player minimizing the non-reachability probability need not have any optimal strategy, even for a BMDP (see [ESY18, Example 3.2]). However, in [ESY18, Theorem 9.4] it was shown that for BMDPs and BSSGs, such player always has a winning strategy in the case when the non-reachability value is 0 (i.e., when the reachability value is 1). But the following example shows that this is not the case for the more general model of BCSGs.

Example 3.1. *The qualitative almost-sure and limit-sure reachability problems for BCSGs do not coincide.*

$$C \xrightarrow{a,\,c} C \qquad\qquad A \xrightarrow{1/2} \emptyset$$

$$C \xrightarrow{a,\,d} C\,A \qquad\qquad A \xrightarrow{1/2} T_{f^*}$$

$$C \xrightarrow{b,\,c} C\,A$$

$$C \xrightarrow{b,\,d} A$$

In the BCSG above, the player minimizing the non-reachability probability has actions a,b in type C, and the other player has actions c,d in type C. The target type is T_{f^*}. We show that, starting at an object of type C, the non-reachability value is 0, but there is no winning strategy for the player minimizing the non-reachability probability (i.e., maximizing the reachability probability) that achieves value exactly 0.

First, construct a corresponding minimax-PPS using Theorem 3.1. For type A, there is a variable x_A and the equation $x_A = 1/2$. For type C, there is a variable x_C and the equation $x_C = Val\left(\begin{bmatrix} x_C & x_C \cdot x_A \\ x_C \cdot x_A & x_A \end{bmatrix}\right)$. Clearly, $x_A^* = 1/2$, $x_C^* = 0$ is a fixed point for the system. To see that it is indeed the GFP (and in fact the only fixed point), if for any $0 < v \leq 1$ we take $x_C = v$, it is not a fixed point. That is because the minimax value of the matrix game $\begin{bmatrix} v & v/2 \\ v/2 & 1/2 \end{bmatrix}$ is strictly less than v.

There is a sequence of static randomized strategies for the player minimizing the non-reachability that achieve non-reachability values arbitrarily close to 0. Namely, for any $\varepsilon > 0$, let strategy τ_ε assign probability $1-\varepsilon$ to action a and probability ε to action b. Fixing strategy τ_ε for the min player, from the minimax-PPS we get a maxPPS with equations $x_A = 1/2$ and $x_C = \max\{x_C \cdot (1-\varepsilon) + x_C \cdot \varepsilon/2, \;\; x_C \cdot (1-\varepsilon)/2 + \varepsilon/2\}$, whose GFP and hence, the optimal non-reachability probabilities vector in the minimax-PPS under strategy τ_ε is $x_A = 1/2$, $x_C = \varepsilon/(1+\varepsilon) \leq \varepsilon$.

However, there is no strategy (static or not) for the min player that achieves non-reachability value exactly 0. To see this, observe that if the min player never puts a positive probability on action b, then the max player can deterministically always choose action c and the game never reaches the target. The very first time that the min player puts any positive probability on action b, then by selecting action d the max player ensures that with a positive probability the game becomes extinct without reaching the target. □

Let us also give a BCSG example that contains types satisfying almost-sure reachability.

Example 3.2. *BCSG example demonstrating almost-sure reachability.*

$$C \xrightarrow{a,\,c} C'\,A \qquad C' \xrightarrow{a',\,c'} C' \qquad A \xrightarrow{1/2} \emptyset$$

$$C \xrightarrow{a,\,d} A \qquad C' \xrightarrow{a',\,d'} A \qquad A \xrightarrow{1/2} T_{f^*}$$

$$C' \xrightarrow{b',\,c'} C'\,A$$

$$C' \xrightarrow{b',\,d'} C'\,A$$

In the BCSG above: the target type is T_{f^*}; the player minimizing the non-reachability probability has actions a', b' in type C' and an action a in type C; and the other player has actions c', d' in type C' and actions c, d in type C. We show that, starting at an object of type C', the non-reachability value is 0 and there is a strategy for the min player that achieves exactly value 0.

Let us again first construct a corresponding minimax-PPS using Theorem 3.1. For type A, there is a variable x_A and the equation $x_A = 1/2$. For type C', there is a variable $x_{C'}$ and the equation $x_{C'} = Val\left(\begin{bmatrix} x_{C'} & x_A \cdot x_{C'} \\ x_A & x_A \cdot x_{C'} \end{bmatrix}\right)$. And for type C, there is a variable x_C and the equation $x_C = Val\left(\begin{bmatrix} x_A \cdot x_{C'} \\ x_A \end{bmatrix}\right) = \max[x_A \cdot x_{C'}, x_A]$. One can check that the GFP of the system is $x_A = x_C = 1/2$, $x_{C'} = 0$.

Furthermore, there is in fact a winning strategy τ for the min player such that, starting at an object of type C', the non-reachability value is 0 (i.e., the reachability value is 1). Namely, at every object of type C', let τ choose deterministically action b'. Then, regardless of the strategy of the max player, with probability 1 infinitely often an independent object of type A will be generated and, hence, infinitely often there will be an independent probability of $1/2$ of hitting the target type. So the overall probability of hitting the target type, starting at an object of type C', is 1.

As for type C, the min player has only one available action and the spoiling strategy for the max player will deterministically select action d. Then, regardless of the min player strategy, it will be guaranteed that the target type is not reached with probability $1/2$. Otherwise, if the max player chooses action c in type C, then an object of type C' will be immediately generated and, therefore, as previously observed the target type will be reached with probability 1, which is in contradiction with the objective of the max player. □

3.2 P-time algorithm for deciding reachability value $= 0$ for BCSGs

In this section we show that there is a P-time algorithm for computing the variables x_i with value $g_i^* = 1$ for the GFP in a given minimax-PPS, or in other words, for a given BCSG, deciding whether the value for reaching the target type, starting with an object of a given type T_i, is 0. The algorithm does not take into consideration the actual probabilities on the transitions in the game (i.e., the coefficients of the polynomials), but rather depends only on the structure of the game (respectively, the dependency graph structure of the minimax-PPS) and performs an AND-OR graph reachability analysis. The algorithm is easy and generalizes the algorithm given for deciding $g_i^* = 1$ for BSSGs in [ESY18, Proposition 4.1].

Algorithm 3.1 Simple P-time algorithm for computing the set of types with reachability value 0 in a given BCSG, or equivalently the set of variables $\{x_i \mid g_i^* = 1\}$ of the associated minimax-PPS.

1. Initialize $S := Z$.
2. Repeat until no change has occurred:
   (a) if there is a variable $x_i \notin S$ of form L or Q such that $P_i(x)$ contains a variable already in S, then add x_i to S.
   (b) if there is a variable $x_i \notin S$ of form M such that for every action $a_{max} \in \Gamma^i_{max}$, there exists an action $a_{min} \in \Gamma^i_{min}$, such that $A_i(x)_{(a_{max},a_{min})} \in S$, then add x_i to S.
3. Output the set $\bar{S} := W - S$.

Proposition 3.3. *Algorithm 3.1 decides, given a BCSG or equivalently a corresponding minimax-PPS, $x = P(x)$, with n variables and GFP $g^* \in [0,1]^n$, for any $i \in [n]$, whether $g_i^* = 1$ or $g_i^* < 1$. Equivalently, for a given BCSG with non-reachability objective and a starting object of type T_i, it decides whether the non-reachability game value is 1. In the case of $g_i^* = 1$, the algorithm produces a deterministic policy (or deterministic static strategy in the BCSG case) σ for the max player (maximizing non-reachability) that forces $g_i^* = 1$. Otherwise, if $g_i^* < 1$, the algorithm produces a mixed policy τ (a mixed static strategy) for the min player (minimizing non-reachability) that guarantees $g_i^* < 1$.*

Proof. Let $W = \{x_1, \ldots, x_n\}$ denote the set of all variables in the minimax-PPS, $x = P(x)$. Recall that the dependency graph of $x = P(x)$ has a directed edge (x_i, x_j) if and only if variable x_i depends on variable x_j, i.e., x_j occurs in $P_i(x)$. Let us call a variable x_i *deficient* if $P_i(x)$ is of form L and $P_i(\mathbf{1}) < 1$. Let $Z \subseteq \{x_1, \ldots, x_n\}$ be the set of deficient variables. The remaining variables $X = W - Z$ are partitioned, according to their SNF-form equations: $X = L \cup Q \cup M$ (refer to Definition 6 for the SNF-form of a minimax-PPS).

The intuition behind the algorithm is as follows: notice that in step 2.(b) no matter what strategy the max player chooses in the particular variable (i.e., type in the game), the min player can ensure with a positive probability to end up in a successor variable that already is bad for the max player. The resulting winning strategies for the players' corresponding winning sets (it is irrelevant to define strategies in the losing nodes) are: (i) for $x_i \in S$, the min player's strategy (mixed static) τ selects uniformly at random among the "witness" moves from step 2.(b), and (ii) for $x_i \in \overline{S}$ the max player's strategy (deterministic static) σ chooses an action $a_{max} \in \Gamma^i_{max}$ that ensures staying within $\overline{S}$ no matter what the minimizer's action is (which must exist, otherwise x_i would have been added to set S).

We need to prove that $g^*_i < 1$ iff $x_i \in S$. First, we show that $x_i \in S$ implies $g^*_i < 1$. Assume $x_i \in S$ (and therefore τ is defined). We analyse by induction, based on the time (iteration) in which variable x_i was added to set S in the iterative algorithm. For the base case, if x_i was added at the initial step (i.e., $x_i \in Z$), then $g^*_i \leq P_i(\mathbf{1}) < 1$. For the induction step, if variable x_i is of form L or Q, then $g^*_i = P_i(g^*)$ is a linear combination (with positive coefficients whose sum is ≤ 1) or a quadratic term, containing at least one variable x_j that was already in set S prior to x_i, and hence, by induction, $g^*_j < 1$. Hence, $g^*_i < 1$. If x_i is of form M, then for $\forall a_{max} \in \Gamma^i_{max}$, $\exists a_{min} \in \Gamma^i_{min}$ such that the corresponding variable $x_{(a_{max},a_{min})} \in S$ (i.e., $g^*_{(a_{max},a_{min})} < 1$), and τ gives positive probability (in fact, probability $\geq \frac{1}{|\Gamma^i_{min}|}$) to all such witnesses a_{min}. For any strategy σ that the maximizer picks, let σ^1 be the part of σ for just the first initial step of the game. In other words, if the game starts in an object of type T_i (variable x_i), then $\sigma^1(x_i)$ denotes the probability distribution on actions Γ^i_{max} that σ assigns in the very first step of the play. Then the reachability probability under the described randomized static strategy τ for the min player and an arbitrary strategy σ for the max player is:

$$\sum_{a_{min},a_{max}} \sigma^1(x_i)(a_{max}) \cdot \tau(x_i)(a_{min}) \cdot (1 - g^*_{(a_{max},a_{min})})$$
$$= \sum_{a_{max}} \sigma^1(x_i)(a_{max}) \sum_{a_{min}} \tau(x_i)(a_{min}) \cdot (1 - g^*_{(a_{max},a_{min})})$$
$$\geq \sum_{a_{max}} \sigma^1(x_i)(a_{max}) \cdot \frac{1}{|\Gamma^i_{min}|} \cdot c = \frac{c}{|\Gamma^i_{min}|}$$

where $c := \min\{1 - g^*_{(a_{max},a_{min})} \mid a_{max} \in \Gamma^i_{max},\ a_{min} \in \Gamma^i_{min} \text{ s.t. } 1 - g^*_{(a_{max},a_{min})} > 0\}$ (note that $c > 0$). It follows that for any strategy $\sigma \in \Psi_1$, $(g^*_{\sigma,\tau})_i \leq 1 - \frac{c}{|\Gamma^i_{min}|}$, or in other words $(g^*_{*,\tau})_i < 1$. Thus, $g^*_i \leq (g^*_{*,\tau})_i < 1$.

Next, to show that if $g^*_i < 1$ then $x_i \in S$, we prove the contrapositive statement. Assume $x_i \in \overline{S}$ (and therefore σ is defined). All variables of form $L \cup Q$ depend only on variables in $\overline{S}$ (otherwise they would have been added to set S). Moreover, for every x_i of form M, there is a maximizer action a_{max} such that, all variables in row a_{max} of the matrix of $A_i(x)$ are in $\overline{S}$. If no such action exists, then x_i would have been added to set S in step 2.(b). Let $\sigma(x_i)$ choose such an action a_{max} deterministically (i.e., with probability 1). In the dependency graph of the resulting (after fixing the defined policy σ) minPPS, $x = P_{\sigma,*}(x)$, there are no edges from $\bar{S}$ to S: all variables of form L, Q, or M depend only on $\bar{S}$ variables, otherwise they would have been added to set S. Moreover, $\overline{S}$ does not contain any deficient variables. So, $P_i(\mathbf{1}) = 1$ for every $x_i \in \overline{S}$, and the all-1 vector is a fixed point for the subsystem of the minPPS, $x = P_{\sigma,*}(x)$, induced by the variables $\bar{S}$. In other words, $(g^*_{\sigma,*})_i = 1$ (thus $g^*_i = 1$) for all $x_i \in \bar{S}$. □

3.3 minimax-PPS preliminary results

Following the definitions introduced in ([ESY18], Section 5), a *linear degenerate (LD)-PPS* is a PPS where every polynomial, $P_i(x)$, is linear and contains no constant term (i.e., $P_i(x) = \sum_{j=1}^{n} p_{ij}x_j$) and where the coefficients $p_{ij} \in [0,1]$ sum to 1. Hence, a LD-PPS has for LFP (q^*) and GFP (g^*) the all-0 and the all-1 vectors, respectively. Furthermore, a PPS that does not contain a linear degenerate bottom strongly-connected component (i.e., a component in the dependency graph that is strongly connected and has no edges going out of it), is called a *linear degenerate free(LDF)-PPS*. In other words, a LDF-PPS is a PPS that satisfies the conditions of Lemma 3.4(ii) below. Given a minimax-PPS, $x = P(x)$, a policy τ for the min player is called LDF if the resulting PPS for all max player policies σ, namely $x = P_{\sigma,\tau}(x)$, is a LDF-PPS. Having introduced this, now we can reference some known results from [ESY18].

Lemma 3.4 (cf. [ESY18], Lemma 5.1)**.** *For any PPS, $x = P(x)$, exactly one of the following two cases holds:*

(i) $x = P(x)$ contains a linear degenerate bottom strongly-connected component (BSCC), S, i.e., $x_S = P_S(x_S)$ is a LD-PPS, and $P_S(x_S) \equiv B_S x_S$, for a stochastic matrix B_S.

(ii) every variable x_i either is, or depends (directly or indirectly) on, a variable x_j where $P_j(x)$ has one of the following properties:

1. $P_j(x)$ has a term of degree 2 or more,

2. $P_j(x)$ has a non-zero constant term, i.e., $P_j(\mathbf{0}) > 0$ or

3. $P_j(\mathbf{1}) < 1$.

Lemma 3.5 (cf. [ESY18], Lemma 5.2)**.** *If a PPS, $x = P(x)$, has either GFP $g^* < \mathbf{1}$, or LFP $q^* > \mathbf{0}$, then $x = P(x)$ is a LDF-PPS.*

Lemma 3.6 (cf. [ESY18], Lemma 5.5)**.** *For any LDF-PPS, $x = P(x)$, and $y < \mathbf{1}$, if $P(y) \leq y$ then $y \geq q^*$ and if $P(y) \geq y$, then $y \leq q^*$. In particular, if $q^* < \mathbf{1}$, then q^* is the only fixed-point q of $x = P(x)$ with $q < \mathbf{1}$.*

The following is a generalized version (for concurrent games) of [ESY18, Lemma 9.1]. In particular, statement (3.) is more involved to prove.

Lemma 3.7. *For a minimax-PPS, $x = P(x)$, if the GFP $g^* < \mathbf{1}$, then:*

1. *there exists a (mixed) LDF policy τ for the min player such that $g^*_{*,\tau} < \mathbf{1}$.*
2. *for any LDF min player's policy τ', it holds that $g^* \leq q^*_{*,\tau'}$.*
3. *there is a sequence of (mixed) LDF policies $(\tau^{(i)})_{i\in\mathbb{N}}$ for the min player such that for every $\varepsilon > 0$, there is $i \geq 0$ where for all $j \geq i$, $\tau^{(j)}$ has the property $g^* \leq q^*_{*,\tau^{(j)}} \leq g^* + \varepsilon$.*

Proof. For point (1.), recall that since $g^* < \mathbf{1}$, the algorithm from the previous Section 3.2 will return a mixed static strategy (policy) τ for the min player such that $g^*_{*,\tau} < \mathbf{1}$. Thus for all max's strategies σ : $g^*_{\sigma,\tau} \leq \sup_{\pi\in\Psi_1} g^*_{\pi,\tau} = g^*_{*,\tau} < \mathbf{1}$. By Lemma 3.5, all PPSs, $x = P_{\sigma,\tau}(x)$, are LDF, which results in the policy τ being LDF as well.

Showing claim (2.), let us fix any LDF policy τ' for the min player. Notice that $g^* = P(g^*) = \inf_\pi P_{*,\pi}(g^*) \leq P_{*,\tau'}(g^*)$. In the resulting maxPPS, there exist a policy σ for the max player such that $g^* \leq P_{*,\tau'}(g^*) = P_{\sigma,\tau'}(g^*)$. For every variable x_i with $g^*_i = \max\{g^*_1, \ldots, g^*_{d_i}\}$ in the maxPPS, the strategy itself chooses the successor in the dependency graph that maximizes g^*_i. Now using Lemma 3.6 with LDF-PPS $x = P_{\sigma,\tau'}(x)$ and $y := g^* < \mathbf{1}$, it follows that $g^* \leq q^*_{\sigma,\tau'} \leq \sup_{\pi\in\Psi_1} q^*_{\pi,\tau'} = q^*_{*,\tau'}$.

Proof of (3.). By statement (1.), there is a mixed LDF policy τ where $q^*_{*,\tau} \leq g^*_{*,\tau} < \mathbf{1}$. Let us start a policy improvement iterative process with $\tau^{(1)} := \tau$. By statement (2.), we know that $g^* \leq q^*_{*,\tau^{(1)}}$ and clearly there exists some $\varepsilon^{(1)} > 0$ such that $q^*_{*,\tau^{(1)}} \leq g^* + \varepsilon^{(1)}$. Suppose that at i-th iteration of the technique, we have a mixed LDF policy $\tau^{(i)}$ with the property $g^* \leq q^*_{*,\tau^{(i)}} \leq g^* + \varepsilon^{(i)}$ (the policy improvement process assumption). If

$P(q^*_{*,\tau^{(i)}}) = q^*_{*,\tau^{(i)}}$, stop the process (we will show that in this case actually $q^*_{*,\tau^{(i)}} = g^*$). Otherwise, since $P(q^*_{*,\tau^{(i)}}) = \inf_{\tau' \in \Psi_2} P_{*,\tau'}(q^*_{*,\tau^{(i)}}) \leq P_{*,\tau^{(i)}}(q^*_{*,\tau^{(i)}}) = q^*_{*,\tau^{(i)}}$, then there is a variable x_j where $P_j(q^*_{*,\tau^{(i)}}) = Val(A_j(q^*_{*,\tau^{(i)}})) < (q^*_{*,\tau^{(i)}})_j$. Note that x_j is indeed of form M, otherwise $P_j(q^*_{*,\tau^{(i)}}) = (P_{*,\tau^{(i)}}(q^*_{*,\tau^{(i)}}))_j$, since for L-form and Q-form variables policy $\tau^{(i)}$ does not have a choice to make. Let the new policy $\tau^{(i+1)}$ be the static strategy that adopts the optimal mixed strategy for the min player in the zero-sum matrix game $A_j(q^*_{*,\tau^{(i)}})$ (exists by the minimax theorem) in variable x_j (type T_j), and stays the same as $\tau^{(i)}$ in all other variables (types). Before moving on with the proof, we will first demonstrate that $\tau^{(i+1)}$ is also LDF.

Claim 3.8. *Policy $\tau^{(i+1)}$ is LDF.*

Proof. Assume $\tau^{(i+1)}$ is not LDF. Then, by Lemma 3.4(i), there is a policy σ for the max player such that in the PPS, $x = P_{\sigma,\tau^{(i+1)}}(x)$, there is a linear degenerate bottom strongly-connected component C. It should contain x_j and all variables that x_j depend directly on in the PPS, i.e., appearing in $(P_{\sigma,\tau^{(i+1)}}(x))_j$. Otherwise C would have also been a linear degenerate BSCC of $x = P_{\sigma,\tau^{(i)}}(x)$ and $\tau^{(i)}$ would not have been LDF.

Due to the construction of the new policy and by standard facts from zero-sum games, we have $P_{*,\tau^{(i+1)}}(q^*_{*,\tau^{(i)}}) \leq q^*_{*,\tau^{(i)}}$ with strict inequality in variable $x_j \in C$, i.e., $(P_{*,\tau^{(i+1)}}(q^*_{*,\tau^{(i)}}))_j \leq Val(A_j(q^*_{*,\tau^{(i)}})) < (q^*_{*,\tau^{(i)}})_j$. Let $j'' = \arg\min_{j' \in C}(q^*_{*,\tau^{(i)}})_{j'}$ be the coordinate in $(q^*_{*,\tau^{(i)}})_C$ with the minimum value. We already know that $(P_{*,\tau^{(i+1)}}(q^*_{*,\tau^{(i)}}))_{j''}$ $\leq (q^*_{*,\tau^{(i)}})_{j''}$.

And we also claim that any $x_{j'} \in C$ satisfies $(q^*_{*,\tau^{(i)}})_{j''} = (q^*_{*,\tau^{(i)}})_{j'}$. That is, in the vector $(q^*_{*,\tau^{(i)}})_C$, any variable $x_{j'} \in C$ has the same minimum value. To show this, consider the form of $(P_{\sigma,\tau^{(i+1)}}(x))_{j''}$. It can not be of Q-form type, due to component C being at the same time *bottom* SCC and linear degenerate in $P_{\sigma,\tau^{(i+1)}}(x)$ (refer to Lemma 3.4). Then it is surely of L-form, and so $(P_{\sigma,\tau^{(i+1)}}(q^*_{*,\tau^{(i)}}))_{j''}$ is a convex combination of some values in $(q^*_{*,\tau^{(i)}})_C$. If any of these values is bigger than the minimum value (namely, $(q^*_{*,\tau^{(i)}})_{j''}$), then $(P_{\sigma,\tau^{(i+1)}}(q^*_{*,\tau^{(i)}}))_{j''} > (q^*_{*,\tau^{(i)}})_{j''}$, which is not true, because $(P_{\sigma,\tau^{(i+1)}}(q^*_{*,\tau^{(i)}}))_{j''} \leq (P_{*,\tau^{(i+1)}}(q^*_{*,\tau^{(i)}}))_{j''} \leq (q^*_{*,\tau^{(i)}})_{j''}$. So for any x_v ($\in C$), appearing in $(P_{\sigma,\tau^{(i+1)}}(x))_{j''}$, we have $(q^*_{*,\tau^{(i)}})_{j''} = (q^*_{*,\tau^{(i)}})_v$, i.e., x_v has the same minimum value. But applying this argument inductively (i.e., for variables appearing in $(P_{\sigma,\tau^{(i+1)}}(x))_v$ and so on) in the closed recurrent set C, we actually get the claim that any $x_{j'} \in C$ satisfies $(q^*_{*,\tau^{(i)}})_{j''} = (q^*_{*,\tau^{(i)}})_{j'}$.

Due to component C being bottom strongly-connected in $x = P_{\sigma,\tau^{(i+1)}}(x)$, it follows that $(q^*_{*,\tau^{(i)}})_j = (q^*_{*,\tau^{(i)}})_{j''}$ and $(q^*_{*,\tau^{(i)}})_k = (q^*_{*,\tau^{(i)}})_{j''}$ for every variable x_k, appearing in

$(P_{\sigma,\tau^{(i+1)}}(x))_j$. Then it means that $(q^*_{*,\tau^{(i)}})_j = (q^*_{*,\tau^{(i)}})_k$ for every such variable x_k, or in other words $(q^*_{*,\tau^{(i)}})_j = (P_{\sigma,\tau^{(i+1)}}(q^*_{*,\tau^{(i)}}))_j$. But we know that $(P_{\sigma,\tau^{(i+1)}}(q^*_{*,\tau^{(i)}}))_j \leq (P_{*,\tau^{(i+1)}}(q^*_{*,\tau^{(i)}}))_j < (q^*_{*,\tau^{(i)}})_j$. This is a contradiction. Therefore, the initial assumption of the claim is false and $\tau^{(i+1)}$ is indeed LDF. □

Going back to the proof of statement (3.) from Lemma 3.7, to recap, $P_{*,\tau^{(i+1)}}(q^*_{*,\tau^{(i)}}) \leq q^*_{*,\tau^{(i)}}$ with strict inequality for variable x_j, because of the construction of $\tau^{(i+1)}$. There is a max player's policy σ such that $q^*_{\sigma,\tau^{(i+1)}} = q^*_{*,\tau^{(i+1)}}$ and $P_{\sigma,\tau^{(i+1)}}(q^*_{*,\tau^{(i)}}) \leq P_{*,\tau^{(i+1)}}(q^*_{*,\tau^{(i)}}) \leq q^*_{*,\tau^{(i)}}$ with strict inequality for variable x_j. Applying Lemma 3.6 to the LDF-PPS, $x = P_{\sigma,\tau^{(i+1)}}(x)$, and $y := q^*_{*,\tau^{(i)}}$, it follows that $q^*_{*,\tau^{(i+1)}} = q^*_{\sigma,\tau^{(i+1)}} \leq q^*_{*,\tau^{(i)}}$. What is more, as $P_{*,\tau^{(i+1)}}(q^*_{*,\tau^{(i)}}) \neq q^*_{*,\tau^{(i)}}$, then it can not be an equality. So the policy improvement algorithm does not visit the same min player policy twice, and $q^*_{*,\tau^{(i+1)}} < q^*_{*,\tau^{(i)}}$. And by the assumption of the policy improvement process, $q^*_{*,\tau^{(i)}} \leq g^* + \varepsilon^{(i)}$. So there exists $0 < \varepsilon^{(i+1)} < \varepsilon^{(i)}$ where $q^*_{*,\tau^{(i+1)}} \leq g^* + \varepsilon^{(i+1)}$. Also by statement (2.) of the Lemma, $g^* \leq q^*_{*,\tau^{(i+1)}}$ and $g^* \leq q^*_{*,\tau^{(i)}}$, since both policies ($\tau^{(i)}$ and $\tau^{(i+1)}$) are LDF.

That shows that the policy improvement process constructs a sequence $(\tau^{(i)})_{i\in\mathbb{N}}$ of mixed LDF policies that bring value $q^*_{*,\tau^{(i)}}$ closer and closer to g^* as $i \to \infty$. So for every $\varepsilon > 0$, there exists $i \geq 0$ such that for all $j \geq i$, $\tau^{(j)}$ has the property $g^* \leq q^*_{*,\tau^{(j)}} \leq g^* + \varepsilon$.

Say that by some chance the process has stopped (say at iteration t) with a mixed LDF policy $\tau^{(t)}$, i.e., $q^*_{*,\tau^{(t)}} = P(q^*_{*,\tau^{(t)}})$ and $q^*_{*,\tau^{(t)}}$ is a fixed point of the minimax-PPS. By statement (2.) of the Lemma, since $\tau^{(t)}$ is LDF, then $g^* \leq q^*_{*,\tau^{(t)}}$. But also because g^* is the GFP, then $g^* \geq q^*_{*,\tau^{(t)}}$. Hence, $g^* = q^*_{*,\tau^{(t)}}$. □

Lemma 3.9 (cf. [ESY18], Lemma 6.1)**.** *For any maxPPS, $x = P(x)$, if GFP $g^* < \mathbf{1}$ then g^* is the unique fixed point of $x = P(x)$ in $[0,1]^n$. In other words, $g^* = q^*$, where q^* is the LFP of $x = P(x)$.*

Lemma 3.10 (cf. [ESY20], Lemma 3.20)**.** *If $\mathbf{0} < q^* < \mathbf{1}$ is the LFP of a max/minPPS, $x = P(x)$, in n variables, then for all $i \in \{1,\dots,n\}$:*

$$1 - q^*_i \geq 2^{-4|P|}$$

In other words, $0 < q^*_i \leq 1 - 2^{-4|P|}$, for all $i \in \{1,\dots,n\}$.

Proposition 3.11. *For a minimax-PPS, $x = P(x)$, with GFP $g^* < \mathbf{1}$, for all $i \in [n]$:*

$$1 - g^*_i \geq 2^{-4|P|}$$

Proof. Since $g^* < \mathbf{1}$, by Lemma 3.7(1. and 2.) there is a LDF policy τ' for the minimizer such that $g^* \leq q^*_{*,\tau'} \leq g^*_{*,\tau'} < \mathbf{1}$. Moreover, fixing τ' in the minimax-PPS, we get a maxPPS, $x = P_{*,\tau'}(x)$, where by Lemma 3.9, there is a unique fixed point $q^*_{*,\tau'} = g^*_{*,\tau'}$. Hence, for all $i \in [n]$, $1 - g^*_i \geq 1 - (g^*_{*,\tau'})_i = 1 - (q^*_{*,\tau'})_i$. And by Lemma 3.10, $1 - (q^*_{*,\tau'})_i \geq 2^{-4|P_{*,\tau'}|} \geq 2^{-4|P|}$, where the last inequality holds due to $|P_{*,\tau'}| \leq |P|$. This is because to encode $x_i = \max\{x_1, x_2, \cdots x_{|\Gamma^i_{max}|}\}$ from maxPPS, $x = P_{*,\tau'}(x)$, there cannot be any more bits needed than to encode $x_i = Val(A_i(x))$ from minimax-PPS, $x = P(x)$, where the dimensions of the matrix are $|\Gamma^i_{max}| \times |\Gamma^i_{min}|$. □

Claim 3.12. *In a LD-PPS, $x = P(x)$:*

1. *in every fixed point, all variables have equal values.*
2. *there are infinitely many fixed points.*

Proof. Recall that in a linear degenerate PPS, $x = P(x)$, every polynomial is linear, with no constant terms, of the form $P_i(x) = \sum_{j=1}^{n} p_{ij} x_j$, where $p_{ij} \in [0,1]$ and $\sum_{j=1}^{n} p_{ij} = 1$ for all $i \in [n]$. Clearly for any $\mathbf{0} \leq x \leq \mathbf{1}$ where all values in x are the same, $P(x) = x$ and so x is a fixed point for the PPS. This shows the second statement.

And in order to show the even stronger statement (1.), take some other fixed point $x \in [0,1]^n$ and let $j'' = \arg\min_{i \in [n]} x_i$ be the index of the variable with the minimum value. Since $P_{j''}(x)$ is a convex combination of some subset of variables $\{x_j \mid j \in [n]\}$, if any of them is larger than the minimum value, then $P_{j''}(x) > x_{j''}$, contradicting that x is a fixed point for $P(\cdot)$. So for any such variable x_j, appearing in $P_{j''}(x)$, $x_j = x_{j''}$. Applying this inductively, the statement follows, since the dependency graph of the PPS, $x = P(x)$, is strongly connected. □

3.4 P-time algorithm for deciding almost-sure reachability for BCSGs

In this section the focus is on the qualitative almost-sure reachability problem for BCSGs, i.e., given a BCSG and starting with an object of a given type T_i, decide whether the reachability value is 1 *and* there exists an optimal strategy to achieve this value for the player aiming to maximize the reachability probability. That is, the algorithm presented here computes a set F of variables (types), such that for any variable $x_i \in F$, starting from one object of corresponding type T_i there is a strategy $\tau \in \Psi_2$ for the player aiming to reach the target type T_{f^*}, such that no matter what the other player does, almost-surely an object of type T_{f^*} will be reached. We of course also wish to compute such a strategy if it exists.

Algorithm 3.2 P-time algorithm for computing the types that satisfy almost-sure reachability in a given BCSG, i.e., the set of variables $\{x_i \mid \exists \tau \in \Psi_2 \ (g^*_{*,\tau})_i = 0\}$ in the associated minimax-PPS.

---

1. Initialize $S := \{x_i \in X \mid P_i(\mathbf{0}) > 0$, that is $P_i(x)$ has a constant term $\}$.
   Let $\gamma^i_0 := \Gamma^i_{min}$ for every variable $x_i \in X - S$. Let $t := 1$.

2. Repeat until no change has occurred to S:

   (a) if there is a variable $x_i \in X - S$ of form L where $P_i(x)$ contains a variable already in S, then add x_i to S.

   (b) if there is a variable $x_i \in X - S$ of form Q where both variables in $P_i(x)$ are already in S, then add x_i to S.

   (c) if there is a variable $x_i \in X - S$ of form M and if for all $a_{min} \in \Gamma^i_{min}$, there exists a $a_{max} \in \Gamma^i_{max}$ such that $A_i(x)_{(a_{max},a_{min})} \in S \cup \{1\}$, then add x_i to S.

3. For each $x_i \in X - S$ of form M, let:
   $\gamma^i_t := \{a_{min} \in \gamma^i_{t-1} \mid \forall a_{max} \in \Gamma^i_{max},\ A_i(x)_{(a_{max},a_{min})} \notin S \cup \{1\}\}$. (Note that $\gamma^i_t \subseteq \gamma^i_{t-1}$.)

4. Let $F := \{x_i \in X - S \mid P_i(\mathbf{1}) < 1$, or $P_i(x)$ is of form Q $\}$

5. Repeat until no change has occurred to F:

   (a) if there is a variable $x_i \in X - (S \cup F)$ of form L where $P_i(x)$ contains a variable already in F, then add x_i to F.

   (b) if there is a variable $x_i \in X - (S \cup F)$ of form M such that for $\forall a_{max} \in \Gamma^i_{max}$, there is a min player's action $a_{min} \in \gamma^i_t$ such that $A_i(x)_{(a_{max},a_{min})} \in F$, then add x_i to F.

6. If $X = S \cup F$, **return** F, and halt.

7. Else, let $S := X - F$, $t := t + 1$, and go to step 2.

We now present the algorithm. First, as a preprocessing step, we apply Algorithm 3.1, which identifies in P-time all the variables x_i where $g^*_i = 1$. We then remove these variables from the system, substituting the value 1 in their place. We then simplify and reduce the resulting SNF-form minimax-PPS into a reduced form, with GFP $g^* < \mathbf{1}$. Note that the resulting reduced SNF-form minimax-PPS may contain some variables x_j

of form M, whose corresponding matrix $A_j(x)$ has some entries that contain the value 1 rather than a variable (because we substituted 1 for removed variables x_j, where $g^*_j = 1$). Note also that in the reduced SNF-form minimax-PPS each variable x_i of form Q has an associated quadratic equation $x_i = x_j x_k$, because if one of the variables (say x_k) on the right-hand side was set to 1 during the preprocessing step, the resulting equation ($x_i = x_j$) would have been declared to have form L in the reduced minimax-PPS. We henceforth assume that the minimax-PPS is in SNF-form, with $g^* < \mathbf{1}$, and we let X be its set of (remaining) variables. We apply now Algorithm 3.2 to the minimax-PPS with $g^* < \mathbf{1}$, which identifies the variables x_i in the minimax-PPS (equivalently, the types in the BCSG), from which we can almost-surely reach the target type T_{f^*} (i.e., $g^*_i = 0$ *and* there is a strategy τ^* for the player minimizing the non-reachability probability that achieves this value, no matter what the other player does).

Theorem 3.13. *Given a BCSG with minimax-PPS, $x = P(x)$, such that the GFP $g^* < \mathbf{1}$, Algorithm 3.2 terminates in polynomial time and returns the following set of variables: $\{x_i \in X \mid \exists \tau \in \Psi_2 \ (g^*_{*,\tau})_i = 0\}$.*

Proof. First, let us provide some notation and terminology for analyzing the algorithm. The integer $t \geq 1$ represents the number of iterations of the main loop of the algorithm, i.e., the number of executions of steps (2.) through (7.) (inclusive; note that some of these steps are themselves loops). Let S_t denote the set S inside iteration t of the algorithm and just before we reach step (3.) of the algorithm (in other words, just after the loop in step (2.) has finished). Similarly, let F_t denote the set F just before step (6.) in iteration t of the algorithm. We also define a new set, K_t, which doesn't appear explicitly in the algorithm. Let $K_t := X - (S_t \cup F_t)$, for every iteration $t \geq 1$. The set γ^i_t in the algorithm denotes a set of moves/actions of the min player at variable x_i (i.e., type T_i). We shall later show that γ^i_t, for $t \geq 1$, is a set of actions such that if the minimizer's strategy only chooses a distribution on actions contained in γ^i_t, for each variable x_i, then starting at any variable $x_j \in X - S_t$, the play will always stay out of S_t.

We now start the proof of correctness for the algorithm. Clearly, the algorithm terminates, i.e., step (6.) eventually gets executed. This is because (due to step (7.)) each extra iteration of the main loop must add at least one variable to the set $S \subseteq X$, and variables are never removed from the set S. It also follows easily that the algorithm runs in P-time, since the main loop executes for at most $|X|$ iterations, and during each such iteration, each nested loop within it also executes at most $|X|$ iterations. So, the proof of correctness requires us to show that when the algorithm halts, the set F is indeed the winning set for the minimizer (i.e., the player that aims to minimize the non-reachability probability). That is, we need to show that for all $x_i \in F$ there exists a (not-necessarily static) strategy τ for the minimizing player such that $(g^*_{*,\tau})_i = 0$, i.e., regardless of what strategy σ the maximizer plays again τ the probability of *not* reaching the target is 0. On the other hand, if $x_i \in S$, we need to show that there is no such strategy τ for the minimizer that forces $(g^*_{*,\tau})_i = 0$. In fact, we will show that for all $x_i \in S$ the following stronger property $(**)_i$ holds:

$(**)_i$: There is a strategy σ for the maximizing player, such that for any strategy τ of the minimizing player $(g^*_{\sigma,\tau})_i > 0$; in other words, starting with one object of type T_i, using strategy pair σ and τ, there is a positive probability of never reaching the target type.

Note that property $(**)_i$ does not rule out that $g^*_i = 0$, because even if $(**)_i$ holds it is possible that $\inf_{\tau\in\Psi_2}(g^*_{\sigma,\tau})_i = 0$. In such a case, it would mean that starting in an object of type T_i, almost-sure reachability cannot be achieved but limit-sure reachability can. That is discussed later in Section 3.5.

First, let us show that if variable $x_i \in S$ when the algorithm terminates, then $(**)_i$ holds.

Lemma 3.14. *For every $x_i \in S$, property $(**)_i$ is satisfied.*

Proof. To show this, we use an induction on the "time" when a variables is added to set S. That is, if all variables x_j added to set S in previous steps and previous iterations satisfy $(**)_j$, then if a new variable x_i is added to set S, it must also satisfy $(**)_i$. In the process of proving this, we shall in fact construct a single non-static randomized strategy σ for the max player that ensures that for all $x_i \in S$, regardless what strategy τ the min player plays against σ, the probability of not reaching the target starting at one object of type T_i is positive.

Consider the initial set S of variables $\{x_i \in X \mid P_i(\mathbf{0}) > 0\}$ that S is initialized to in step (1.) of the algorithm. Clearly all these variables satisfy $g^*_i \geq P_i(\mathbf{0}) > 0$. Thus, for these variables assertion $(**)_i$ holds using *any* strategy σ for the maximizer. Next consider a variable x_i added to set S inside the loop in step (2.) of the algorithm, during some iteration.

(i) If $x_i = P_i(x)$ is of form L, then $P_i(x)$ contains a variable x_j (with a positive coefficient), that was added previously to set S, and hence $(**)_j$ holds. Thus there is a positive probability that one object of type T_i will produce one object of type T_j in the next generation. It thus follows that $(**)_i$ holds, by using the same strategy $\sigma \in \Psi_1$ that witnesses the fact that $(**)_j$ holds.

(ii) If $x_i = P_i(x)$ is of form Q (i.e., $x_i = x_j \cdot x_r$), then $P_i(x)$ has both variables already added to set S, i.e., $(**)_j$ and $(**)_r$ both hold. Then $(**)_i$ also holds, because starting from any object of type T_i, the next generation necessarily contains one object of type T_j and one object of type T_r, and thus by combining the two witness strategies for $(**)_j$ and $(**)_r$, we have a strategy $\sigma \in \Psi_1$ that, starting from one object of type T_i, will ensure a positive probability of not reaching the target, regardless of the strategy $\tau \in \Psi_2$ of the minimizer.

(iii) If $x_i = P_i(x)$ is of form M, then $\forall a_{min} \in \Gamma^i_{min}$, $\exists a_{max} \in \Gamma^i_{max}$ such that $A_i(x)_{(a_{max}, a_{min})} \in S \cup \{1\}$. In this case, let us define the strategy σ to behave as follows at any object of type T_i. For each $a_{min} \in \Gamma^i_{min}$, we designate one "witness" $a_{max}[a_{min}] \in \Gamma^i_{max}$, which witnesses that $A_i(x)_{(a_{max}[a_{min}], a_{min})} \in S \cup \{1\}$. Then, at any object of type T_i, σ chooses uniformly at random among the witnesses $a_{max}[a_{min}]$ for all $a_{min} \in \Gamma^i_{min}$. So, starting with one object of type T_i, no matter what strategy the min player chooses, there is a positive probability that in the next step that object will either not produce any offspring (in the case where $A_i(x)_{(a_{max}[a_{min}], a_{min})} = 1$) and hence not reach the target, or else will generate a single successor object of a type T_j, associated with variable $x_j = A_i(x)_{(a_{max}[a_{min}], a_{min})}$ that already belongs to set S, and hence such that $(**)_j$ holds. Hence, by combining with the strategies that witness such $(**)_j$ with the local (static) behavior of σ described for any object of type T_i, we obtain a strategy σ that witnesses the fact that $(**)_i$ holds.

Now consider any variable x_i that is added to set S in step (7.) of some iteration t, in other words any variable $x_i \in K_t := X - (S_t \cup F_t)$. Since all variables in set K_t were not added to sets S_t or F_t during iteration t, we must have that: (A.) x_i satisfies $P_i(\mathbf{1}) = 1$ and $P_i(\mathbf{0}) = 0$; (B.) x_i is not of form Q; (C.) if x_i is of form L, then it depends directly only on variables in K_t; and (D.) if x_i is of form M, then:

$$\exists a_{max} \in \Gamma^i_{max} \text{ such that } \forall a_{min} \in \gamma^i_t,\ A_i(x)_{(a_{max}, a_{min})} \notin (F_t \cup S_t \cup \{1\}). \tag{3.2}$$

Let $(q_h)_{h=0}^{\infty}$, $h \in \mathbb{N}$ be the infinite sequence of increasing probabilities defined by: $q_h = 2^{-(1/2^h)}$. Note that as $h \to \infty$, the probability q_h approaches 1 from below.

Given a finite history H of height h (meaning the depth of the forest that the history represents is h), for any object e in the current generation (the leaves) of H, if the object e has type T_i such that the associated variable $x_i \in K_t$ is of form M, we shall construct the strategy σ to behave as follows starting at the object e. The strategy σ will choose one action a_{max} that "witnesses" the statement (3.2) above, and will place probability q_h on that action, and it will distribute the remaining probability $1 - q_h$ uniformly among all actions in Γ^i_{max}. We claim that this strategy σ ensures that for any object e of type T_i such that $x_i \in K_t$, irrespective of the strategy of the minimizing player, the probability of not reaching the target type T_{f^*} starting with e (at any point in history) is positive. This clearly implies that $\forall \tau \in \Psi_2 : (g^*_{\sigma,\tau})_{K_t} > \mathbf{0}$. To prove this, there are two cases here:

1. First, suppose that during the entire play of the game, at all objects e whose type T_i such that $x_i \in K_t$ has form M, the min player only uses actions belonging to γ^i_t. Then in the resulting history of play there *can not* be any such object e who does not generate a child or whose child in the history (a necessarily unique child, since e has form M) is an object e' of a type (variable) in set S_t (this is because step (3.) of the algorithm, which defines γ^i_t, ensures that actions for the min player in γ^i_t can not possibly produce a child in set S_t or no child at all, no matter what the max player does). Furthermore, such an object e, occurring at depth h in history, must with positive probability $\geq q_h$, produce a child e' with a type in K_t (because of point (D.) above, and because of the fact that the max player plays at e a witness a_{max} to the statement (3.2) with probability $\geq q_h$).

   So consider an object e of some type (variable) in set K_t, that occurs in a history H at height $h \geq 0$, and consider the tree of descendants of e. Recall that by point (B.) above, this tree of descendants does not contain objects of Q-form types. What is the probability, under the strategy σ, and under any strategy τ for the min player whose moves are confined to the sets specified by γ_t, that the "tree" of descendants of e is just a "line" consisting of an infinite sequence of objects $e_0 = e$, e_1, e_2, $\ldots$, all of which have types (variables) contained in set K_t? This probability is clearly at least

$$\prod_{d=h}^{\infty} q_d = \prod_{d=h}^{\infty} 2^{-(1/2^d)} \geq \prod_{d=0}^{\infty} 2^{-(1/2^d)} = 2^{-\sum_{d=0}^{\infty}(1/2^d)} = 2^{-2} = \frac{1}{4}$$

   That is, irrespective of what strategy τ is played by the minimizer, there is a positive probability bounded away from 0 (indeed, $\geq 1/4$) of staying forever confined in objects having types (variables) in set K_t. In such a case, clearly, there will be positive probability of not reaching the target type (since the types (variables) in set K_t are not the target type).

2. Next suppose that, on the other hand, there is a history H of some height h and a leaf e of H that has type T_i where M-form $x_i \in K_t$, such that the min player's strategy τ plays at object e some action(s) outside of the set γ_t^i with a positive probability. Note that for all actions $a'_{min} \notin \gamma_t^i$, there is a max player's action $a_{max} \in \Gamma^i_{max}$ such that $A_i(x)_{(a_{max},a'_{min})} \in S_t \cup \{1\}$. Note moreover that the strategy σ assigns positive probability, at least $(1-q_h)/|\Gamma^i_{max}|$ to every action in $\Gamma^i_{\max}$. Thus, if the min player's strategy τ puts a positive probability $\tau(H,e,a_{min}) > 0$ on some action $a_{min} \notin \gamma_i^t$, then with probability $\geq \left(\max_{a_{min} \notin \gamma_t^i} \tau(H,e,a_{min})\right) \cdot \frac{(1-q_h)}{|\Gamma^i_{max}|}$, either the object e will have no child (since we can have $A_i(x)_{(a_{max},a_{min})} = 1$), or the only child of object e in the history will be an object e' whose type is in the set S_t, from which we already know that the target type T_{f^*} is *not* reached with a positive probability. So in either case, with a positive probability the target type $T^*_{f^*}$ will not be reached from descendants of e.

Now, let us assume the max player uses this strategy σ, and suppose we start play at one object e' of type T_i such that $x_i \in K_t$. Suppose, first, that during the entire history of play the min player's strategy τ uses only actions in γ_t^i for all variables $x_i \in K_t$ of form M. In this case, with a positive probability bounded away from 0 (in fact $\geq 1/4$), the play tree after k rounds (i.e., depth k), for any positive $k \geq 1$, consists of simply a linear sequence of objects having types (variables) in set K_t. Thus in this case, with probability $\geq 1/4$, the play will forever stay in set K_t, and will never reach target type T_{f^*}. On the other hand, suppose the min player's strategy τ does at some point in some history consisting entirely of a linear sequence of objects of types (variables) in set K_t, namely at some specific object e of type in set K_t at depth h, plays an action outside of γ_t^i with a positive probability. Then σ ensures that with a positive probability (albeit a probability depending on h and thus not bounded away from 0) either e will have no child or the unique child of e will be an object of type T_j such that $x_j \in S_t$, i.e., there is a positive probability of *not* reaching the target T_{f^*} from the descendants of e, and thus also from the start of the game (because we assumed the play staring from e' and up to e consists of a linear sequence of objects all having types (variables) in set K_t). Thus, for all strategies $\tau \in \Psi_2$, and all $x_i \in K_t$, $(g^*_{\sigma,\tau})_i > 0$. Note however, that in general it may be the case that $\inf_\tau (g^*_{\sigma,\tau})_i = 0$, because in the case when τ does play outside of γ_t^i, the probability of not hitting the target type is not bounded away from 0 (it depends both on the depth h at which τ first moves outside of γ_t^i with a positive probability, and it also depends on the probability of that move, and for both reasons it can be arbitrarily close to 0). This establishes the first part of the proof of Theorem

3.13, i.e., that for every $x_i \in S$ the property $(**)_i$ holds. □

Now we proceed to the second part of the proof. Suppose F is the set of variables output by the algorithm when it halts (and that therefore $S = X - F$). Suppose the algorithm executed exactly t^* iterations of the main loop before halting (so that the value of t just before halting is t^*). We will show that there is a (randomized non-static) strategy τ of the minimizing player such that, for all $x_i \in F$, regardless what strategy σ the maximizer employs, starting with an object of type T_i, the probability of *not* reaching the target type is 0. In other words, that $(g^*_{*,\tau})_i = 0$, which is what we want to prove.

Lemma 3.15. *There is a randomized non-static strategy $\tau \in \Psi_2$ such that, for every $x_i \in F$, $(g^*_{*,\tau})_i = 0$.*

Proof. Before describing τ, we first describe a static randomized strategy (i.e., a mixed policy) τ^* for the minimizing player, that will eventually lead us toward a definition of τ.

Specifically, we define the mixed policy (randomized static strategy) τ^* as follows. Let τ' be any LDF policy such that $g^*_{*,\tau'} < \mathbf{1}$. Such an LDF policy τ' must exist, by Lemma 3.7(1.). For all variables $x_i \in S$, let $\tau^*(x_i) := \tau'(x_i)$. In other words, at all variables $x_i \in S$, let τ^* behave according to the exact same distribution on actions as the LDF policy τ'. For every variable $x_i \in F$ of form M, define τ^* as follows: note that x_i must have entered set F in some iteration of the inner loop in step (5.)(b) of the algorithm, during the final iteration t^* of the main loop. Therefore, for all $a_{max} \in \Gamma^i_{max}$, there exists a "witness" action $a_{min}[a_{max}] \in \gamma^i_{t^*}$ such that the associated variable $A_i(x)_{(a_{max},a_{min}[a_{max}])}$ was already in set F, before x_i was added to set F. For $x_i \in F$ we define the policy τ^* at variable x_i, i.e., the distribution $\tau^*(x_i)$, to be the uniform distribution over the set $\{a_{min}[a_{max}] \in \gamma^i_{t^*} \mid a_{max} \in \Gamma^i_{max}\}$ of such "witnesses".

We now wish to show that τ^*, as defined, is itself an LDF policy. Consider any fixed policy (i.e., static randomized strategy) σ for the max player, and consider the resulting system of polynomial equations $x = P_{\sigma,\tau^*}(x)$. For every variable $x_i \in F$, consider the variables x_i depends on directly in the equation $x_i = (P_{\sigma,\tau^*}(x))_i$. Let's consider separately the cases, based on the form of equation $x_i = P_i(x)$: (1) if $x_i = P_i(x)$ is of form L, then in $x_i = (P_{\sigma,\tau^*}(x))_i$ the variable x_i depends directly only on variables in set F, because otherwise it would have been added to set S; (2) if x_i is of form M, then again it depends directly only on variables in set F, because $\tau^*(x_i)$ only puts positive probability on actions in $\gamma^i_{t^*}$; (3) if x_i is of form Q, then x_i depends directly on at least one variable in set F, because otherwise it would have been added to set S. Since there is a clear order in which variables were added to set F and due to the initialization of F (step (4.)), in the dependency graph of $x = P_{\sigma,\tau^*}(x)$ every variable in F satisfies one of the three conditions in Lemma 3.4(ii) (namely, 1. or 3.). So for every variable $x_i \in X$, consider the paths in the dependency graph of $x = P_{\sigma,\tau^*}(x)$ starting at x_i:

- either there exists a path from x_i in this dependency graph to variable $x_j \in F$, which in turn must have a path to a variable $x_{j'}$ such that either $P_{j'}(\mathbf{1}) < 1$, or $x_{j'}$ has form Q. In either case, this means that x_i satisfies one of the conditions of Lemma 3.4(ii) (namely, either condition (1.) or condition (3.)); Or
- all paths from x_i only contain variables in set S. But for all variables $x_k \in S$, $\tau^*(x_k)$ is exactly the same distribution as $\tau'(x_k)$, and since the LDF policy τ' was chosen so that $g^*_{*,\tau'} < \mathbf{1}$, this means that there is a path from x_i to a variable x_j satisfying one of the three conditions in Lemma 3.4(ii) (specifically, condition (3.)).

Therefore, $x = P_{\sigma,\tau^*}(x)$ is a LDF-PPS. But since the fixed policy σ was arbitrary, this implies that strategy τ^* is indeed an LDF policy. Since τ^* is LDF, by Lemma 3.7(2.), it holds that $g^* \leq q^*_{*,\tau^*}$.

We now construct a *non-static* strategy τ, which combines the behavior of the two policies (i.e., two static strategies) τ' and τ^* in a suitable way, such that for all $x_i \in F$, $(g^*_{*,\tau})_i = 0$. In other words, τ will be a strategy for the minimizer such that, no matter what strategy σ the maximizer uses starting with one object of type T_i, the probability of not reaching the target type is 0.

The non-static strategy τ is defined as follows. The strategy τ will, in each generation, declare one object in the current generation to be the "queen" (and this object will always have a type (variable) in set F). Other objects in each generation will be "workers". Assume play starts at a single object e of some type T_i such that $x_i \in F$. We declare this object the "queen" in the initial population. If the queen e has associated variable x_i of form M, then τ plays at e according to distribution $\tau^*(x_i)$. This results, (with probability 1), regardless of the strategy of the maximizer, in some successor object e' in the next generation of type T_j such that $x_j \in F$. In this case, we declare e' the queen in the next generation, and we apply the same strategy τ starting at the queen e' of the next generation, as if the game is starting at this single object e' of type T_j. If the variable x_i associated with the queen e is of form L, then in the next generation either we hit the target (with probability $(1 - P_i(\mathbf{1}))$, or (with probability $P_i(\mathbf{1})$) we generate a single successor object e' of some type T_j such that $x_j \in F$. In this latter case again, we declare e' the queen of the next generation, and we use the same strategy τ that is being defined, and apply it to e' as if the game is starting with the single object e'. If the queen e has associated variable x_i of form Q, then in the next generation there are two successor objects, e' and e'' of types T_j and T_k respectively (these may be the same type), such that either $x_j \in F$ or $x_k \in F$, or both are in set F. In this case, we choose one of the two successors whose type (variable) is in set F, say w.l.o.g. that this is e', and we declare e' the queen of the next generation, we proceed from e' using the same strategy τ that is being defined, as if the game starts with the single object e'. However, we declare the other object e'' a "worker", and starting with e'' and thereafter (in the entire subtree of play rooted at e'') we use the static strategy (i.e., the LDF policy) τ'. This completes the definition of the non-static strategy τ.

We now show that indeed τ satisfies that, no matter what strategy σ the maximizer uses against it, for any $x_i \in F$, starting with one object of type T_i, the probability of not reaching the target type is 0. In other words, we show that using τ the probability of reaching the target type is 1, no matter what the opponent does.

To see this, first note that the LDF policy τ' was chosen so that $g^*_{*,\tau'} < \mathbf{1}$. Thus, since in the resulting max-PPS, $x = P_{*,\tau'}(x)$, the player maximizing non-reachability probability always has a static optimal strategy (by Theorem 3.1), it follows that the subtree of the play rooted at any "worker" object e'' starting at which strategy τ' is applied by the min player, has a positive probability $(1 - g^*_{*,\tau'})_i > 0$ (in fact, $\geq 2^{-4|P|}$ by Proposition 3.11) of eventually reaching the target type.

Next note that the sequence of queens is finite if and only if we have hit the target. Next, we establish that if the sequence of queens is infinite, then, with probability 1, infinitely often the queen is of form Q and thus in the next generation it generates both a queen and a worker. Thus, because of the infinite sequence of workers generated by queens, there will be infinitely many independent chances of hitting the target with probability at least $\min_i(1 - g^*_{*,\tau'})_i$ (in fact, $\geq 2^{-4|P|}$). Hence, we will hit the target (somewhere in the entire tree of play) with probability 1.

It remains to show that, if the sequence of queens is infinite, then, with probability 1, infinitely often a queen is of form Q. We in fact claim that with a positive probability bounded away from 0, in the next $n = |X|$ generations either we reach a queen of form Q, or the queen has the target as a child. To see this, we note that each variable $x_i \in F$ has entered set F in some iteration of the loop in step (5.) of the algorithm (in the last iteration of the main loop). We can thus define inductively, for each variable $x_i \in F$, a finite tree R_i, rooted at x_i, which shows "why" x_i was added to set F. Specifically, if $P_i(\mathbf{1}) < 1$ or x_i has form Q, then R_i consists of just a single node (leaf) labeled by x_i. If x_i has form L, then it was added in step (5.) because $P_i(x)$ has a variable x_j that was already in set F. In this case, the tree R_i has an edge from the root, labeled by x_i, to a single child labeled by x_j, such that this child is the root of a subtree R_j. If x_i has form M then R_i has a root labeled by x_i and has a child labeled by variable $x_j = A_i(x)_{(a_{max}, a_{min}[a_{max}])} \in F$ and has R_j as a subtree, for each $a_{max} \in \Gamma^i_{max}$ and where $a_{min}[a_{max}] \in \gamma^i_{t^*}$ is the "witness" for a_{max}, in the condition that allows step 5.(b) of the algorithm to add x_i to set F.

Clearly the tree R_i is finite and has depth at most n (since there are only n variables, and there is a strict order in which the variables entered the set F).

Now we argue that starting at a queen of type T_i, using strategy τ for the minimizing player, with a positive probability bounded away from 0 in the next n steps the sequence of queens will follow a root-to-leaf path in R_i, regardless of the strategy of the max player. To see this, note that if a node is labeled by x_j is of form L, then the play will in the next step, with probability associated with the transition in the BCSG move to the unique child (the new queen) $x_{j'}$ that is the immediate child of the root in R_j, and thus next will be at the root of the subtree $R_{j'}$. If the node is labeled by x_j of form M, then irrespective of the distribution on actions played by the max player, in the next step with a positive probability bounded away from 0, we will move to a child $x_{j'} = A_i(x)_{(a_{max}, a_{min}[a_{max}])} \in F$ which is a child of the root in R_j, itself rooted at a subtree $R_{j'}$, because at queen objects we are using policy τ^* for the minimizer. Thus, starting at a queen x_i, with a positive probability bounded away from 0, within n steps the play arrives a leaf of the tree R_i. If the leaf corresponds to a variable x_j with $P_j(\mathbf{1}) < 1$, then the process will reach in the next step the target type with a positive probability bounded away from 0. If, on the other hand, the leaf corresponds to a variable x_j of form Q, then the queen generates two children. The probability that the queen reaches infinitely often a leaf of form L with $P_j(\mathbf{1}) < 1$ but does not reach the target is 0. Thus, if the queen never reaches the target throughout the play, then the queen will generate more than one child infinitely often with probability 1, and hence will generate infinitely many independent workers with probability 1. By the choice of the policy τ' followed by workers, the subtree rooted at each worker will hit the target with a positive probability bounded away from 0. Hence, the probability of hitting the target type is 1. □

This completes the proof of Theorem 3.13. □

Corollary 3.16. *Let F be the set of variables output by Algorithm 3.2.*

1. *Let $S := X - F$. There is a randomized non-static strategy σ for the max player (maximizing non-reachability) such that for all $x_i \in S$, and for all strategies τ of the min player (minimizing non-reachability), starting with one object of type T_i, the probability of reaching the target type is < 1.*

2. *There is a randomized non-static strategy τ for the min player (minimizing non-reachability), such that for all strategies σ of the max player (maximizing non-reachability), and for all $x_i \in F$, starting at one object of type T_i the probability of reaching the target type is 1.*

Proof. 1. The strategy σ constructed in the proof of Theorem 3.13 for all variables $x_i \in S$ achieves precisely this.

2. The strategy τ constructed in the proof of Theorem 3.13 for all variables $x_i \in F$ achieves precisely this. □

Remark: Both the strategy σ from Corollary 3.16(1) and the strategy τ from 3.16(2) are *non-static* strategies. However, we note that both of these non-static randomized strategies have suitable compact descriptions (as functions that map finite histories to distributions over actions for objects in the current populations), and that both these strategies can be constructed and described compactly in polynomial time, as a function of the encoding size of the input BCSG.[3]

3.5 P-time algorithm for deciding limit-sure reachability for BCSGs

In this section, we focus on the qualitative limit-sure reachability problem for BCSGs, i.e., given a BCSG and starting with one object of a given type T_i, decide whether the reachability value is 1. Recall that there may not exist an optimal strategy for the player aiming to reach the target T_{f^*}, which was the question in the previous section (almost-sure reachability). However, there may nevertheless be a sequence of strategies that achieve values arbitrarily close to 1 (limit-sure reachability), and the question of the

[3]However, it is worth pointing out that the functions that these strategies compute, i.e., functions from histories to distributions, need not themselves be polynomial-time as a function of the encoding size of the history: this is because the probabilities on actions that are involved can be double-exponentially small (and double-exponentially close to 1), as a function of the size of the history.

existence of such a sequence is what we address in this section. Since we translate reachability into non-reachability when analysing the corresponding minimax-PPS, we are asking whether there exists a sequence of strategies $\langle \tau^*_{\varepsilon_j} \mid j \in \mathbb{N} \rangle$ for the min player, such that $\forall j \in \mathbb{N}$, $\varepsilon_j > \varepsilon_{j+1} > 0$, and where $\lim_{j \to \infty} \varepsilon_j = 0$, such that the strategy $\tau^*_{\varepsilon_j}$ forces the non-reachability probability to be at most ε_j, regardless of the strategy σ used by the max player. In other words, for a given starting object of type T_i, we ask whether $\inf_{\tau \in \Psi_2} (g^*_{*,\tau})_i = 0$.

Again, as in the almost-sure case, we first, as a preprocessing step, use the P-time algorithm from Proposition 3.3 (i.e., Algorithm 3.1) to remove all variables x_i such that $g^*_i = 1$, and we substitute 1 for these variables in the remaining equations. We hence obtain a reduced SNF-form minimax-PPS, for which we can assume $g^* < \mathbf{1}$. The set of all remaining variables in the SNF-form minimax-PPS is again denoted by X. Thereafter, we apply Algorithm 3.3, which computes the set of variables, x_i, such that $g^*_i = 0$. In other words, we compute the set of types, such that starting from one object of that type the value of the reachability game is 1. Before considering Algorithm 3.3 in detail, we provide some preliminary results that will be used to prove its correctness. More precisely, we first examine the nested loop in step (4.)(b) of the algorithm. This inner loop is derived directly from a closely related "limit-escape" construction used by de Alfaro, Henzinger, and Kupferman in [dAHK07] (see Algorithm 4 and section 4.4.2 in the cited paper). Proofs are provided here for the facts needed about this construction.

3.5.1 Limit-escape

For a variable x_i of form M, for 1-step local strategies $\sigma(x_i)$ and $\tau(x_i)$ at x_i for the two players (i.e., $\sigma(x_i)$ and $\tau(x_i)$ are distributions on Γ^i_{max} and Γ^i_{min}, respectively), and for a set $W \subseteq X \cup \{1\}$ which can include both variables and possibly also the constant 1, let us define:

$$p(x_i \to W, \sigma(x_i), \tau(x_i)) = \sum_{\{(a_{max}, a_{min}) \in \Gamma^i_{max} \times \Gamma^i_{min} | A_i(x)_{(a_{max}, a_{min})} \in W\}} \sigma(x_i)(a_{max}) \cdot \tau(x_i)(a_{min})$$

Thus $p(x_i \to W, \sigma(x_i), \tau(x_i))$ denotes the probability that, starting with one object of type T_i, and using the 1-step strategies specified by $\sigma(x_i)$ and $\tau(x_i)$, we will either generate a child object of type T_j such that $x_j \in W$, or (only if $1 \in W$) generate no child object (i.e., go extinct in the next generation).

Algorithm 3.3 P-time algorithm for computing the types that satisfy limit-sure reachability in a given BCSG, i.e., the set of variables $\{x_i \mid g_i^* = 0\}$ in the associated minimax-PPS.

1. Initialize $S := \{x_i \in X \mid P_i(\mathbf{0}) > 0$, that is $P_i(x)$ has a constant term $\}$.

2. Repeat until no change has occurred to S:

   (a) if there is a variable $x_i \in X - S$ of form L where $P_i(x)$ contains a variable already in S, then add x_i to S.

   (b) if there is a variable $x_i \in X - S$ of form Q where both variables in $P_i(x)$ are already in S, then add x_i to S.

   (c) if there is a variable $x_i \in X - S$ of form M and if for all $a_{min} \in \Gamma^i_{min}$, there exists $a_{max} \in \Gamma^i_{max}$ such that $A_i(x)_{(a_{max},a_{min})} \in S \cup \{1\}$, then add x_i to S.

3. Let $F := \{x_i \in X - S \mid P_i(\mathbf{1}) < 1$, or $P_i(x)$ is of form Q $\}$

4. Repeat until no change has occurred to F:

   (a) if there is a variable $x_i \in X - (S \cup F)$ of form L where $P_i(x)$ contains a variable already in F, then add x_i to F.

   (b) if there is a variable $x_i \in X - (S \cup F)$ of form M and if the following procedure returns "Yes", then add x_i to F.

      i. Set $L_0 := \emptyset$, $B_0 := \emptyset$, $k := 0$. Let $O := X - (S \cup F)$.

      ii. Repeat:

         - $k := k + 1$.
         - $L_k := \{a_{min} \in \Gamma^i_{min} - \bigcup_{j=0}^{k-1} L_j \mid \forall a_{max} \in \Gamma^i_{max} - B_{k-1},\ A_i(x)_{(a_{max},a_{min})} \in F \cup O\}$.
         - $B_k := B_{k-1} \cup \{a_{max} \in \Gamma^i_{max} - B_{k-1} \mid \exists a_{min} \in L_k\ s.t.\ A_i(x)_{(a_{max},a_{min})} \in F\}$.

         Until $B_k = B_{k-1}$.

      iii. Return: "Yes" if $B_k = \Gamma^i_{max}$, and "No" otherwise.

5. If $X = S \cup F$, **return** F, and halt.

6. Else, let $S := X - F$, and go to step 2.

Informally, the following is the high-level intuition behind the limit-escape technique. Recall from the almost-sure algorithm section that when we were computing the set of variables (types) that almost-surely reach the target (starting with one object of the type), we were looking for those ones that the min player can force in the next round with a positive probability to stay within set F without any risk of immediate extinction or entering set S (since entering set S would mean a positive probability of *not* reaching the target type). That is, informally, in the almost-sure algorithm section, we were keeping track of variables x_j (types T_j) for which there exists an 1-step local strategy $\tau'(x_j)$ for the min player such that:

$$\inf_{\sigma(x_j)\in\mathcal{D}(\Gamma^j_{max})} p(x_j \to F, \sigma(x_j), \tau'(x_j)) > 0$$

$$\sup_{\sigma(x_j)\in\mathcal{D}(\Gamma^j_{max})} p(x_j \to S\cup\{1\}, \sigma(x_j), \tau'(x_j)) = 0$$

where $\mathcal{D}(\Gamma^j_{max})$ denotes the set of distributions on the set of actions Γ^j_{max}. However, in the limit-sure reachability case, the aim is to reach the target type with probability arbitrarily close to 1. So a small chance to enter set $S\cup\{1\}$ can be permitted, as long as the ratio of the 1-step probability between entering set F and set $S\cup\{1\}$ can be made arbitrarily high. That is, in addition to the aforementioned variables x_j, we also want to keep track of variables x_i for which: regardless of the min player's 1-step local strategy, there is a positive probability to enter set $S\cup\{1\}$ in the next step; but there is a family of 1-step local strategies $\tau_e(x_i)$, $e\to 0$, for the min player such that:

$$\lim_{e\to 0} \inf_{\sigma(x_i)\in\mathcal{D}(\Gamma^i_{max})} \frac{p(x_i\to F, \sigma(x_i), \tau_e(x_i))}{p(x_i\to S\cup\{1\}, \sigma(x_i), \tau_e(x_i))} = \infty$$

More formally, consider step (4.)(b) of Algorithm 3.3 and assume that for a variable x_i the answer is "Yes". Let $N := \max_i |\Gamma^i_{min}|$. Given some $0\le e\le \frac{1}{2N}$, consider the following static distribution, $safe(x_i,e)$, on actions for the min player at x_i (i.e., distribution on Γ^i_{min}):

$$safe(x_i,e)(a_{min}) := \begin{cases} (e^2)^{j-1}\cdot\dfrac{(1-e^2)}{|L_j|} & \text{if } a_{min}\in L_j,\text{ for some } j\in\{1,\dots,k-1\} \\[2ex] (e^2)^{k-1}\cdot\dfrac{1}{|\Gamma^i_{min}-\bigcup_{q=0}^{k-1}L_q|} & \text{otherwise} \end{cases} \tag{3.3}$$

Lemma 3.17. *Suppose that for a variable $x_i\in X-(S\cup F)$ the answer in step (4.)(b) of the algorithm is "Yes", and for any e such that $0\le e\le\frac{1}{2N}$, let $\tau_e(x_i) = safe(x_i,e)$. Then for every 1-step local strategy (i.e., distribution on actions in Γ^i_{max}), $\sigma(x_i)$, for the* *max player, the following inequality holds:*

$$p(x_i \to F, \sigma(x_i), \tau_e(x_i)) > \frac{1}{e} \cdot p(x_i \to S \cup \{1\}, \sigma(x_i), \tau_e(x_i))$$

Proof. Since the answer is "Yes" for x_i, when the loop in step (4.)(b) stops at some iteration m, we must have $B_m = B_{m-1} = \Gamma^i_{max}$. Suppose $\sigma(x_i)$ is any 1-step local strategy for the max player at x_i. Let q_j denote the probability that the max player distributes among all its actions in the set $B_j - B_{j-1}$. Since $B_{m-1} = \Gamma^i_{max}$, then clearly $\sum_{j=1}^{m-1} q_j = 1$.

Since each action $a \in \Gamma^i_{max}$ was added at some point to set $B_m = B_{m-1}$, there is some $1 \leq j_a \leq m-1$, such that there exists $a'_{min} \in L_{j_a}$, such that $A_i(x)_{(a,a'_{min})} \in F$. Moreover, $\tau_e(x_i)(a'_{min}) = \left(e^2\right)^{j_a - 1} \cdot \frac{\left(1-e^2\right)}{|L_{j_a}|}$. And furthermore, we know from the definitions of the L and B sets, that for all $a_{min} \in \bigcup_{q=0}^{j_a-1} L_q$, $A_i(x)_{(a,a_{min})} \in O$. With this information, we can give bounds on the 1-step probabilities of visiting sets F and $S \cup \{1\}$ under $\tau_e(x_i)$ and $\sigma(x_i)$:

$$p(x_i \to F, \sigma(x_i), \tau_e(x_i)) \geq \sum_{j=1}^{m-1} q_j \cdot \left(e^2\right)^{j-1} \cdot \frac{\left(1-e^2\right)}{|L_{j_a}|} \geq \sum_{j=1}^{m-1} q_j \cdot \left(e^2\right)^{j-1} \cdot \frac{\left(1-e^2\right)}{N}$$

$$p(x_i \to S \cup \{1\}, \sigma(x_i), \tau_e(x_i)) \leq \sum_{j=1}^{m-1} q_j \cdot \left(e^2\right)^{j}$$

The second inequality in the first row follows from the fact that $L_{j_a} \subseteq \Gamma^i_{min}$ (i.e., $|L_{j_a}| \leq |\Gamma^i_{min}| \leq N$). The inequality in the second row follows from the fact that the maximum probability of ending up in set $S \cup \{1\}$ in the next round occurs when for all maximizer actions a_{max} in each segment $B_j - B_{j-1}$, all remaining minimizer actions $a_{min} \in \Gamma^i_{min} - \bigcup_{q=0}^{j} L_q$ satisfy $A_i(x)_{(a_{max},a_{min})} \in S \cup \{1\}$. To all these minimizer actions, the distribution $\tau_e(x_i)$ assigns a total probability of $\left(e^2\right)^j$.

Then in order to prove the inequality from the Lemma we need to show that:

$$\sum_{j=1}^{m-1} q_j \cdot \left(e^2\right)^{j-1} \cdot \frac{\left(1-e^2\right)}{N} > \frac{1}{e} \cdot \sum_{j=1}^{m-1} q_j \cdot \left(e^2\right)^{j}$$

First, note that for all $1 \leq j \leq m-1$, since $0 \leq e \leq \frac{1}{2N}$, then $e^2 < \frac{1}{2}$ and we have:

$$\frac{q_j \cdot \left(e^2\right)^{j-1} \cdot \frac{\left(1-e^2\right)}{N}}{q_j \cdot \left(e^2\right)^{j}} = \frac{1-e^2}{e^2 \cdot N} \geq \frac{1-e^2}{(e/2)} > \frac{1}{e}$$

Thus, for all $1 \leq j \leq m-1$, $q_j \cdot \left(e^2\right)^{j-1} \cdot \dfrac{(1-e^2)}{N} > \dfrac{1}{e} \cdot q_j \cdot \left(e^2\right)^j$. And summing over all $1 \leq j \leq m-1$, we get what we wanted to prove. □

Notice that, as a consequence to this Lemma, there is a sequence of 1-step local strategies for the min player such that the ratio of the 1-step probability between entering sets F and $S \cup \{1\}$ diverges over the limit as $e \to 0$.

Assume the opposite, that in step (4.)(b) for a variable x_i the loop stops at some iteration m (i.e., $B_{m-1} = B_m$), but $B_m \overset{\subset}{\neq} \Gamma^i_{max}$, and hence step (4.)(b) answers "No", and x_i is not added to set F. In such a case, let us define the following 1-step local strategy, $\sigma(x_i)$ for the max player which will be used in the next lemma. Let $D^i_{max} := \Gamma^i_{max} - B_m$. Let

$$\sigma(x_i)(a_{max}) := \begin{cases} \dfrac{1}{|D^i_{max}|} & \text{for every } a_{max} \in D^i_{max} \\ 0 & \text{otherwise} \end{cases} \tag{3.4}$$

Lemma 3.18. *Suppose that for a variable $x_i \in X - (S \cup F)$ the answer in step (4.)(b) of the algorithm is "No", and let $\sigma(x_i)$ be defined as in (3.4). Then, there is a constant $c_i > 0$ such that for every 1-step local strategy $\tau(x_i)$ for the min player at x_i, the following inequality holds:*

$$p(x_i \to S \cup \{1\}, \sigma(x_i), \tau(x_i)) \geq c_i * p(x_i \to (F \cup S \cup \{1\}), \sigma(x_i), \tau(x_i))$$

Proof. Suppose the loop from step (4.)(b) stops at iteration m, such that $B_{m-1} = B_m \subset \Gamma^i_{max}$. There are two possibilities:

1. $L_m = \emptyset$: That is, for every $a_{min} \in \Gamma^i_{min} - \bigcup_{q=0}^{m-1} L_q$, there exists $a_{max} \in D^i_{max} = \Gamma^i_{max} - B_{m-1}$ such that $A_i(x)_{(a_{max}, a_{min})} \in S \cup \{1\}$. Let $\tau(x_i)$ be an arbitrary 1-step local strategy for the min player and let $\sigma(x_i)$ be as defined in (3.4). Also let $D^i_{min} := \Gamma^i_{min} - \bigcup_{q=0}^{m-1} L_q$. Then it follows that:

$$\begin{aligned} p(x_i \to S \cup \{1\}, \sigma(x_i), \tau(x_i)) &\geq \sum_{a_{min} \in D^i_{min}} \frac{1}{|D^i_{max}|} \tau(x_i)(a_{min}) \\ &= \frac{1}{|D^i_{max}|} \sum_{a_{min} \in D^i_{min}} \tau(x_i)(a_{min}) \end{aligned} \tag{3.5}$$

Note that, by construction, for all $a_{max} \in D^i_{max}$ and $a_{min} \in \bigcup_{q=0}^{m-1} L_q$, $A_i(x)_{(a_{max},a_{min})} \in O$. Hence, since the support of distribution $\sigma(x_i)$ is D^i_{max}, and since $D^i_{min} = \Gamma^i_{min} - \bigcup_{q=0}^{m-1} L_q$, we have:

$$p(x_i \to (F \cup S \cup \{1\}), \sigma(x_i), \tau(x_i)) \leq \sum_{a_{min} \in D^i_{min}} \tau(x_i)(a_{min}) \tag{3.6}$$

Combining these bounds, we get:

$$\begin{aligned} p(x_i \to S \cup \{1\}, \sigma(x_i), \tau(x_i)) &\geq \frac{1}{|D^i_{max}|} \sum_{a_{min} \in D^i_{min}} \tau(x_i)(a_{min}) \\ &\geq \frac{1}{|D^i_{max}|} p(x_i \to (F \cup S \cup \{1\}), \sigma(x_i), \tau(x_i)) \end{aligned}$$

2. $L_m \neq \emptyset$, but $\{a_{max} \in \Gamma^i_{max} - B_{m-1} \mid \exists a_{min} \in L_m \; s.t. \; A_i(x)_{(a_{max},a_{min})} \in F\} = \emptyset$. Therefore for all $a_{max} \in \Gamma^i_{max} - B_{m-1} = \Gamma^i_{max} - B_m = D^i_{max}$, and for all $a_{min} \in L_m$, $A_i(x)_{(a_{max},a_{min})} \in O$. Let $\tau(x_i)$ be any 1-step local strategy for the min player, and let $\sigma(x_i)$ be as defined in (3.4). Let $D^i_{min} := \Gamma^i_{min} - \bigcup_{q=0}^{m} L_q$. Note that if $D^i_{min} = \emptyset$, then $p(x_i \to S \cup \{1\}, \sigma(x_i), \tau(x_i)) = 0 = p(x_i \to (F \cup S \cup \{1\}), \sigma(x_i), \tau(x_i))$, since support for $\sigma(x_i)$ is D^i_{max} and, by construction, for all $a_{max} \in D^i_{max}$ and $a_{min} \in \Gamma^i_{min} - D^i_{min}$, $A_i(x)_{(a_{max},a_{min})} \in O$. So, in this case, the lemma holds for any constant $c_i > 0$. If $D^i_{min} \neq \emptyset$, then both the inequalities (3.5) and (3.6) hold again, with the minor modification that now we have $D^i_{min} = \Gamma^i_{min} - \bigcup_{q=0}^{m} L_q$ instead of $D^i_{min} = \Gamma^i_{min} - \bigcup_{q=0}^{m-1} L_q$.

Therefore, in both cases the lemma is satisfied with $c_i := \frac{1}{|D^i_{max}|} = \frac{1}{|\Gamma^i_{max} - B_m|}$. □

3.5.2 Limit-sure algorithm

We are now ready to prove correctness for Algorithm 3.3.

Theorem 3.19. *Given a BCSG with minimax-PPS, $x = P(x)$, with GFP $g^* < \mathbf{1}$, Algorithm 3.3 terminates in polynomial time, and returns the set of variables $\{x_i \in X \mid g^*_i = 0\}$.*

Proof. The fact that the algorithm terminates and runs in polynomial time is again evident, as in case of the almost-sure algorithm. (The only new fact to note is that the new inner loop in step (4.)(b), can iterate at most $\max_i |\Gamma^i_{max}|$ times because with each new iteration, k, at least one action is added to the set B_{k-1}, or else the algorithm halts.)

We need to show that when the algorithm terminates, for all $x_i \in F$, $g^*_i = 0$, and for all $x_i \in S = X - F$, $g^*_i > 0$.

Let us first show that for all $x \in S$, $g^*_i > 0$. In fact, we will show that there is a strategy $\sigma \in \Psi_1$, and a vector $b > \mathbf{0}$ of values, such that for all $x_i \in S$, $(g^*_{\sigma,*})_i \geq b_i > 0$.

Lemma 3.20. *There is a strategy $\sigma \in \Psi_1$ and a vector $b > \mathbf{0}$ such that, for every $x_i \in S$, $(g^*_{\sigma,*})_i \geq b_i > 0$.*

Proof. We use an induction for this proof. For the base case, since any variable x_i contained in set S at the initialization step has $g^*_i \geq P_i(\mathbf{0}) > 0$, we have $(g^*_{\sigma,*})_i > P_i(\mathbf{0}) > 0$ for *any* strategy $\sigma \in \Psi_1$, so let $b_i := P_i(\mathbf{0})$. For the inductive step, first consider any variable x_i added to set S in step (2.), in some iteration of the main loop of the algorithm.

(i) If $x_i = P_i(x)$ is of form L, then $P_i(x)$ has a variable x_j already in set S, and by induction $(g^*_{\sigma,*})_j \geq b_j > 0$. Since $P_i(x)$ is linear, with a term $p_{ij} \cdot x_j$, such that $p_{ij} > 0$, we see that $(g^*_{\sigma,*})_i \geq p_{ij} \cdot b_j > 0$, so let $b_i := p_{ij} \cdot b_j$.

(ii) If $x_i = P_i(x)$ is of form Q (i.e., $x_i = x_j \cdot x_r$), then $P_i(x)$ has both variables previously added to set S, i.e., $(g^*_{\sigma,*})_j \geq b_j > 0$ and $(g^*_{\sigma,*})_r \geq b_r > 0$. Then clearly $(g^*_{\sigma,*})_i \geq b_j \cdot b_r > 0$. So let $b_i := b_j \cdot b_r$.

(iii) If $x_i = P_i(x)$ is of form M, then $\forall a_{min} \in \Gamma^i_{min}$, $\exists a_{max} \in \Gamma^i_{max}$ such that $A_i(x)_{(a_{max},a_{min})} \in S \cup \{1\}$. For each $a_{min} \in \Gamma^i_{min}$, let us use $a_{max}[a_{min}] \in \Gamma^i_{max}$, to denote a "witness" to this fact, i.e., such that $A_i(x)_{(a_{max}[a_{min}],a_{min})} \in S \cup \{1\}$. Let strategy σ do as follows: in any object of type T_i corresponding to x_i, σ selects uniformly at random an action from the set $\{a_{max}[a_{min}] \in \Gamma^i_{max} \mid a_{min} \in \Gamma^i_{min}\}$ of all such witnesses. Clearly then, for any $a_{min} \in \Gamma^i_{min}$, the probability that σ at an object of type T_i will choose the witness action $a_{max}[a_{min}]$ is at least $\frac{1}{|\Gamma^i_{max}|}$ (and in fact is also at least $\frac{1}{|\Gamma^i_{min}|}$). So, using σ, starting with one object of type T_i, no matter what strategy the min player chooses, there is a positive probability $\geq \frac{1}{|\Gamma^i_{max}|}$ that either the object will have no child or the object will generate a single child object of a type T_j, associated with variable $x_j = A_i(x)_{(a_{max},a_{min})} \in S$, and hence such that $(g^*_{\sigma,*})_j \geq b_j > 0$. So no matter what strategy the min player picks, there is at least $\frac{1}{|\Gamma^i_{max}|}$ probability that the unique child object belongs to set S, or that there is no child object. Hence, $(g^*_{\sigma,*})_i \geq \frac{1}{|\Gamma^i_{max}|} \cdot \min\{b_j \mid x_j \in S\} > 0$, and again we let $b_i := \frac{1}{|\Gamma^i_{max}|} \cdot \min\{b_j \mid x_j \in S\}$.

Now consider any variable x_i added to set S in step (6.) at some iteration of the algorithm (i.e., $x_i \in K := X - (S \cup F)$). Because x_i was not previously added to sets S or F, then:

(A.) x_i satisfies $P_i(\mathbf{0}) = 0$ and $P_i(\mathbf{1}) = 1$;

(B.) x_i is not of form Q;

(C.) if x_i is of form L, then it depends directly only on variables in set K; and

(D.) if x_i is of form M, then the answer for x_i in step (4.)(b) (during the latest iteration of the main loop) was "No".

For each $x_i \in K$ of form M, let $\sigma(x_i)$ be a probability distribution on actions in Γ^i_{max} defined in (3.4). Let strategy σ use the 1-step local strategy $\sigma(x_i)$ at every object of type T_i encountered during history. We show that, for every $x_i \in K$, $(g^*_{\sigma,*})_i \geq b_i$ for some $b_i > 0$.

By Lemma 3.18, for each variable $x_i \in K$ of form M, and for any arbitrary 1-step local strategy $\tau(x_i)$ for the min player at x_i, there exists $c_i > 0$ such that:

$$p(x_i \rightarrow S \cup \{1\}, \sigma(x_i), \tau(x_i)) \geq c_i * p(x_i \rightarrow (F \cup S \cup \{1\}), \sigma(x_i), \tau(x_i))$$

For $r \geq 1$, let $Pr^{\sigma,\tau}_{x_i}(K \bigcup_{=r} (S \cup \{1\}))$ denote the probability that, starting with one object of type T_i, where $x_i \in K$, using strategy σ as defined above and an arbitrary (not necessarily static) strategy τ, the history of play will stay in the set K for $r-1$ rounds, and in the r-th will either transition to an object whose type is in the set S, or will die (i.e., produce no children). Define $Pr^{\sigma,\tau}_{x_i}(K \bigcup_{=r} (F \cup S \cup \{1\}))$ similarly. The following claim is a simple corollary of Lemma 3.18. Let $c := \min\{c_i \mid x_i \in K\}$. (Note that $0 < c \leq 1$.)

Claim 3.21. *For any integer $r \geq 1$, and for any (not necessarily static) strategy τ for the min player, $Pr^{\sigma,\tau}_{x_i}(K \bigcup_{=r} (S \cup \{1\})) \geq c * Pr^{\sigma,\tau}_{x_i}(K \bigcup_{=r} (F \cup S \cup \{1\}))$.*

Proof. Let $H(x_i, K, r-1)$ denote the set of all sequences of types (variables) in set K of length $r-1$, starting with a type corresponding to variable $x_i \in K$. Recall by point (B.) above that there are no Q-form variables (types) in set K. For a history (sequence) $h \in H(x_i, K, r-1)$, let $l(h)$ denote the index of the variable associated with the last type in h, i.e., the one occurring at round $r-1$. For each $h \in H(x_i, K, r-1)$ there is some probability $q_h \geq 0$ that, starting at an object of type corresponding to variable $x_i \in K$, the population follows the history h for $r-1$ rounds. So

$$
\begin{aligned}
Pr_{x_i}^{\sigma,\tau}(K \mathsf{U}_{=r}(S \cup \{1\})) &= \sum_{h \in H(x_i,K,r-1)} q_h \cdot p(x_{l(h)} \to S \cup \{1\}, \sigma(x_{l(h)}), \tau(x_{l(h)})) \\
&\geq \sum_{h \in H(x_i,K,r-1)} q_h \cdot c_{l(h)} \cdot p(x_{l(h)} \to (F \cup S \cup \{1\}), \sigma(x_{l(h)}), \tau(x_{l(h)})) \\
&\geq c \cdot \sum_{h \in H(x_i,K,r-1)} q_h \cdot p(x_{l(h)} \to (F \cup S \cup \{1\}), \sigma(x_{l(h)}), \tau(x_{l(h)})) \\
&= c \cdot Pr_{x_i}^{\sigma,\tau}(K \mathsf{U}_{=r}(F \cup S \cup \{1\}))
\end{aligned}
$$

where the first inequality follows from Lemma 3.18. □

We now argue that for all $x_i \in K$, there exists $b_i > 0$ such that for any strategy τ for the min player, $(g^*_{\sigma,\tau})_i \geq b_i > 0$.

Consider any strategy τ for the min player. For $x_i \in K$, let $Pr_{x_i}^{\sigma,\tau}(\Box K)$ denote the probability that the history stays forever in set K, starting at one object of type T_i. Let $Pr_{x_i}^{\sigma,\tau}(K \mathsf{U}(S \cup \{1\}))$ denote the probability that the history stays in set K until it eventually either dies (has no children) or transitions to an object with type in set S. Note that:

$$
\begin{aligned}
(g^*_{\sigma,\tau})_i &\geq Pr_{x_i}^{\sigma,\tau}(\Box K) + Pr_{x_i}^{\sigma,\tau}(K \mathsf{U}(S \cup \{1\})) \cdot \min\{(g^*_{\sigma,*})_j \mid x_j \in S\} \\
&\geq Pr_{x_i}^{\sigma,\tau}(\Box K) + Pr_{x_i}^{\sigma,\tau}(K \mathsf{U}(S \cup \{1\})) \cdot \min\{b_j \mid x_j \in S\}
\end{aligned}
$$

We will show that, regardless of the strategy τ for the min player, this probability must be at least:

$$
b_i := \frac{c}{2} \cdot \min\{b_j \mid x_j \in S\}
$$

where $c := \min\{c_i \mid x_i \in K\}$. Recall that $0 < c \leq 1$. Let $p = Pr_{x_i}^{\sigma,\tau}(\Box K)$. If $p \geq \frac{c}{2}$, then we are done, since the inequalities above imply $(g^*_{\sigma,\tau})_i \geq \frac{c}{2} \geq b_i$. So, suppose $p < \frac{c}{2}$. Observe that:

$$
\begin{aligned}
Pr_{x_i}^{\sigma,\tau}(K \mathsf{U}(S \cup \{1\})) &= Pr_{x_i}^{\sigma,\tau}((K \mathsf{U}(S \cup \{1\})) \cap \neg \Box K) \\
&= Pr_{x_i}^{\sigma,\tau}(K \mathsf{U}(S \cup \{1\}) \mid \neg \Box K) \cdot Pr_{x_i}^{\sigma,\tau}(\neg \Box K) \\
&= Pr_{x_i}^{\sigma,\tau}(K \mathsf{U}(S \cup \{1\}) \mid \neg \Box K) \cdot (1-p) \\
&\geq Pr_{x_i}^{\sigma,\tau}(K \mathsf{U}(S \cup \{1\}) \mid \neg \Box K) \cdot \frac{1}{2}
\end{aligned}
$$

So it only remains to show that $Pr_{x_i}^{\sigma,\tau}(K \mathsf{U}(S \cup \{1\}) \mid \neg \Box K) \geq c$. Note that the event $\neg \Box K$ is equivalent to the event $K \mathsf{U}(F \cup S \cup \{1\})$. The event $K \mathsf{U}(S \cup \{1\})$ is equivalent to the disjoint union $\bigcup_{r=1}^{\infty} K \,\mathsf{U}_{=r} (S \cup \{1\})$. Likewise for the event $K \,\mathsf{U} (F \cup S \cup \{1\})$. Therefore:

$$\begin{aligned} Pr_{x_i}^{\sigma,\tau}(K \,\mathsf{U} (S \cup \{1\}) \mid \neg\Box K) &= \frac{Pr_{x_i}^{\sigma,\tau}(K \,\mathsf{U} (S \cup \{1\}))}{Pr_{x_i}^{\sigma,\tau}(\neg\Box K)} \\ &= \frac{\sum_{r=1}^{\infty} Pr_{x_i}^{\sigma,\tau}(K \,\mathsf{U}_{=r} (S \cup \{1\}))}{\sum_{r=1}^{\infty} Pr_{x_i}^{\sigma,\tau}(K \,\mathsf{U}_{=r} (F \cup S \cup \{1\}))} \end{aligned} \tag{3.7}$$

But by Claim 3.21, for all $r \geq 1$, $Pr_{x_i}^{\sigma,\tau}(K \,\mathsf{U}_{=r} (S \cup \{1\})) \geq c \cdot Pr_{x_i}^{\sigma,\tau}(K \,\mathsf{U}_{=r} (F \cup S \cup \{1\}))$. Hence, summing over all r, we have $\sum_{r=1}^{\infty} Pr_{x_i}^{\sigma,\tau}(K \,\mathsf{U}_{=r} (S \cup \{1\})) \geq c \cdot \sum_{r=1}^{\infty} Pr_{x_i}^{\sigma,\tau}(K \,\mathsf{U}_{=r} (F \cup S \cup \{1\}))$. Hence, dividing out (note that the division is well-defined as we have assumed a positive probability for eventually exiting K, i.e., $Pr_{x_i}^{\sigma,\tau}(\neg\Box K) = 1 - p \geq \frac{1}{2}$) and using (3.7), we have $Pr_{x_i}^{\sigma,\tau}(K \,\mathsf{U} (S \cup \{1\}) \mid \neg\Box K) \geq c$.

Thus, $(g^*_{\sigma,\tau})_i \geq b_i$, and since this holds for an arbitrary strategy τ for the min player, we have $(g^*_{\sigma,*})_i \geq b_i > 0$. □

We next want to show that if F is the set of variables output by the algorithm when it halts, then for all variables $x_i \in F$, $g^*_i = 0$, or in other words, that the following holds:

$$\forall \varepsilon > 0,\ \exists \tau_\varepsilon \in \Psi_2 \ \ s.t.\ \forall \sigma \in \Psi_1,\ \ (g^*_{\sigma,\tau_\varepsilon})_i \leq \varepsilon \tag{3.8}$$

Let $N := \max_i |\Gamma^i_{min}|$. Given some $0 \leq e \leq \frac{1}{2N}$, recall from (3.3) the static distribution, $safe(x_i, e)$, on actions Γ^i_{min} for the min player at x_i.

Given an $\varepsilon > 0$, we define a (static) strategy τ_ε as follows. If a variable x_i of form M is in set S, then we let $\tau_\varepsilon(x_i)$ be the uniform distribution on the corresponding action set Γ^i_{min}. For variables in set F, we define τ_ε as follows. Consider the last execution of the main loop of the algorithm. Let $F_0 = \{x_i \in X - S \mid P_i(\mathbf{1}) < 1$, or $P_i(x)$ is of form Q $\}$ be the set of variables assigned to set F in step (3.), and let $x_{i_1}, x_{i_2}, \ldots, x_{i_{k^*}}$ be the variables in set $F - F_0$ ordered according to the time at which they were added to set F in the iterations of step (4.). For each variable $x_{i_t} \in F$ of form M we let $\tau_\varepsilon(x_{i_t}) = safe(x_{i_t}, e_t)$ where the parameters e_t are set as follows. Let n be the number of variables, and $N := \max_i |\Gamma^i_{min}|$ be the maximum number of actions of the min player for any variable of form M. Let κ be the minimum of (1) $1/N$, (2) the minimum (non-negative) coefficient of a monomial in $P_i(x)$ over all variables x_i of form L, and (3) the minimum of $1 - P_i(\mathbf{1})$ over all x_i of form L such that $P_i(\mathbf{1}) < 1$. Let $\lambda = \kappa^n$. Clearly, λ is a rational number that depends on the given minimax-PPS $x = P(x)$ (and the corresponding BCSG) and it has polynomial number of bits in the size of P. Let $d_0 = \lceil \log(\frac{n}{\varepsilon\lambda}) \rceil$ and let $d_t = d_0 \cdot (2N)^t$ for $t \geq 1$. We set $e_t = 2^{-d_t}$ for all $t \geq 0$. The numbers e_t can be doubly-exponentially small, but they can be represented compactly in floating point, i.e., in polynomial size in the size of P and of ε. Note from the definitions that $e_0 \leq \varepsilon\lambda/n$, and $e_t = (e_{t-1})^{2N}$ for all $t \geq 1$.

Consider the maxPPS, $x = P_{*,\tau_\varepsilon}(x)$, obtained from the given minimax-PPS, $x = P(x)$, by fixing the strategy of the min player to policy τ_ε. For every variable x_i of form L or Q, the corresponding equation $x_i = P_i(x)$ stays the same, and for every variable x_i of form M the equation becomes $x_i = \max_{a_{max} \in \Gamma^i_{max}} \{\sum_{a_{min} \in \Gamma^i_{min}} \tau_\varepsilon(x_i)(a_{min}) \cdot A_i(x)_{(a_{max},a_{min})}\}$. Let $f^* = g^*_{*,\tau_\varepsilon}$ be the greatest fixed point of the maxPPS $x = P_{*,\tau_\varepsilon}(x)$, and let $M = \max\{f^*_i | x_i \in F\}$. We will show that $M \leq \varepsilon$, i.e., $f^*_i \leq \varepsilon$ for all $x_i \in F$.

First, we show that all variables of X have value strictly less than 1 in f^*, and we also bound the value of the variables of set S in terms of M.

Claim 3.22.
(1) For all $x_i \in X$, $f^*_i < 1$.
(2) For all $x_i \in X$, $f^*_i \leq \lambda M + (1-\lambda)$.

Proof. The algorithm of Proposition 3.3 (see Algorithm 3.1) computes the set X of variables x_i of the minimax-PPS such that $g^*_i < 1$ (this set is denoted S in Algorithm 3.1, but to avoid confusion with the set S of the limit-sure reachability Algorithm 3.3, we refer to it as X in the following). It is the same set X as the one used in Algorithm 3.3. We use induction on the time that a variable x_i was added to set X in Algorithm 3.1 to show the claim. For part (2), our induction hypothesis is that if a variable x_i is added to set X at time t (where the initialization is time 1) then $f^*_i \leq \kappa^t M + (1-\kappa^t)$. This inequality implies (2) since $t \leq n$ and $\lambda = \kappa^n$.

For the basis case ($t = 1$), x_i is a deficient variable, i.e. $P_i(\mathbf{1}) < 1$, hence $f^*_i \leq P_i(\mathbf{1}) \leq 1-\kappa < 1$.

For the induction step, if x_i is of form L or Q, then $P_i(x)$ contains a variable x_j that was added earlier to set X, hence $f^*_i < 1$ follows from $f^*_j < 1$ by the induction hypothesis. For part (2), if x_i is of form L, then the coefficient of x_j in $P_i(x)$ is at least κ and $f^*_j \leq \kappa^{t-1}M + (1-\kappa^{t-1})$ by the induction hypothesis, hence $f^*_i \leq \kappa(\kappa^{t-1}M + (1-\kappa^{t-1})) + 1 - \kappa = \kappa^t M + (1-\kappa^t)$. If x_i is of form Q, then $f^*_i \leq f^*_j \leq \kappa^{t-1}M + (1-\kappa^{t-1}) \leq \kappa^t M + (1-\kappa^t)$.

If x_i is of form M then for every action $a_{max} \in \Gamma^i_{max}$, there exists an action $a_{min} \in \Gamma^i_{min}$ such that the variable $x_j = A_i(x)_{(a_{max},a_{min})}$ was added previously to set X, and hence its value in f^* is < 1 by the induction hypothesis. Since $\tau_\varepsilon(x_i)$ plays all the actions of Γ^i_{min} with nonzero probability, both when $x_i \in S$ and when $x_i \in F$, it follows that $f_i^* < 1$. This shows part (1). For part (2), if $x_i \in F$, then $f_i^* \leq M \leq \kappa^t M + (1-\kappa^t)$, where the first inequality follows from the definition of M. Suppose $x_i \in S$ and let a_{max} be an action in Γ^i_{max} that yields the greatest fixed point f_i^* in the maxPPS equation $x_i = (P_{*,\tau_\varepsilon}(x))_i$. The right-hand side for this action is a linear expression that contains a variable $x_j = A_i(x)_{(a_{max},a_{min})}$ that was added previously to set X, and the coefficient of this term is $1/|\Gamma^i_{min}| \geq 1/N \geq \kappa$, since $\tau_\varepsilon(x_i)$ is the uniform distribution for $x_i \in S$. Therefore, $f_i^* \leq \kappa f_j^* + (1-\kappa) \leq \kappa(\kappa^{t-1}M + (1-\kappa^{t-1})) + 1 - \kappa = \kappa^t M + (1-\kappa^t)$. □

We can show the key lemma now.

Lemma 3.23. *For all $x_i \in F$, $f_i^* \leq \varepsilon$.*

Proof. Recall that $F = F_0 \cup \{x_{i_1}, x_{i_2}, \ldots, x_{i_{k^*}}\}$. Let $M_0 = \max\{f_i^* | x_i \in F_0\}$ and let $M_t = f_{i_t}^*$ for $t \geq 1$ be the value of x_{i_t} in the greatest fixed point f^* of the maxPPS, $x = P_{*,\tau_\varepsilon}(x)$. Thus, $M = \max\{M_t \mid t \geq 0\}$. Let $r_t = (e_t)^{2N-1}$. Note that for every $x_{i_t} \in F$ of form M, the probability with which $\tau_\varepsilon(x_{i_t}) = safe(x_{i_t}, e_t)$ plays any action in a set L_j is at least $(e_t^2)^{N-1}(1-e_t^2)/N$ which is $> (e_t)^{2N-1} = r_t$ because $e_t < 1/(2N)$. Let $s_t = \Pi^t_{j=1} r_j$; by convention, $s_0 = 1$.

We will show first that for all $t \geq 0$, there exist $a_t, g_t \geq 0$ that satisfy $a_t \geq \lambda \cdot s_t$ and $g_t \leq t \cdot e_0 \cdot a_t/\lambda$, and such that $M_t \leq a_t M^2 + (1 - a_t - g_t)M + g_t$. We will use induction on t.

Basis: $t = 0$. Then $M_0 = f_i^*$ for a variable $x_i \in F_0$ which is either a deficient variable of form L or a variable of form Q. If x_i is of form L, then note that (1) P_i does not contain a constant term (because otherwise x_i would have been added to set S in step (1.)), (2) all the variables of $P_i(x)$ are not in set S (because otherwise x_i would have been added to set S in step (2.)), hence they are all eventually added to set F and thus their value in f^* is at most M, and (3) the coefficients sum to at most $1-\kappa$ because $P_i(\mathbf{1}) < 1$. Therefore, $M_0 = f_i^* \leq (1-\kappa)M \leq \lambda M^2 + (1-\lambda)M$. If x_i is of form Q, at least one of the variables of $P_i(x)$ must belong to set F (because otherwise x_i would have been added to set S in step (2.)), hence its value in f^* is at most M, and the value of the other variable is at most $\lambda M + (1-\lambda)$ by Claim 3.22. Therefore, $M_0 = f_i^* \leq M(\lambda M + 1 - \lambda) = \lambda M^2 + (1-\lambda)M$. Thus in both cases, $M_0 \leq \lambda M^2 + (1-\lambda)M$. We can take $a_0 = \lambda$, $g_0 = 0$.

Induction step: We have $M_t = f_{i_t}^*$. If x_{i_t} is of form L, then $P_{i_t}(x)$ contains a variable x_j that was added earlier to set F; its coefficient, say p, is at least κ. Note again that $P_{i_t}(x)$ does not contain a constant term, all the other variables of $P_{i_t}(x)$ are not in set S, hence they are all eventually added to set F and their value in f^* is at most M, and the sum of their coefficients is $1-p$. Since the variable x_j was added earlier to set F, by the induction hypothesis we have $f_j^* \leq a_u M^2 + (1-a_u-g_u)M + g_u$ for some $u \leq t-1$. Therefore, $M_t \leq p(a_u M^2 + (1-a_u-g_u)M + g_u) + (1-p)M = a_t M^2 + (1-a_t-g_t)M + g_t$, with $a_t = pa_u$ and $g_t = pg_u$. Since $u \leq t-1$, we have $a_u \geq \lambda \cdot s_u \geq \lambda \cdot s_{t-1}$, and since $p \geq \kappa \geq r_t$ it follows that $a_t = pa_u \geq \lambda \cdot s_{t-1} \cdot r_t = \lambda \cdot s_t$. Also, $g_t = pg_u \leq pue_0 a_u/\lambda \leq te_0 a_t/\lambda$.

Suppose x_{i_t} is of form M, and let $a_{max} \in \Gamma^{i_t}_{max}$ be an action of the max player that yields the greatest fixed point $f^*_{i_t}$ in the maxPPS equation $x_{i_t} = (P_{*,\tau_\varepsilon}(x))_{i_t}$. Then a_{max} belongs to some B_j in step (4.) of Algorithm 3.3, and thus there is a $a_{min} \in L_j$ such that the variable $A_{i_t}(x)_{(a_{max},a_{min})}$ was added earlier to set F, i.e., it is variable x_{i_u} for some $u \leq t-1$ or it belongs to F_0. The probability $p = \tau_\varepsilon(x_{i_t})(a_{min})$ of this action in strategy τ_ε is $p = (e_t^2)^{j-1} \cdot (1-e_t^2)/|L_j|$. All the variables $A_{i_t}(x)_{(a_{max},a)}$ for $a \in \cup_{q=1}^j L_q$ are not in set S, hence they are all eventually assigned to set F. The total probability that strategy τ_ε gives to the actions $a \in \cup_{q=1}^j L_q$ is $1-(e_t^2)^j$, hence the remaining probability assigned to the other actions $a \in \Gamma^{i_t}_{min} - \cup_{q=1}^j L_q$ is $(e_t^2)^j$ which is $\leq pe_t$ since $e_t \leq 1/(2N)$. Therefore, $M_t \leq pM_u + (1-p-pe_t)M + pe_t$ for some $u \leq t-1$. By the induction hypothesis, $M_u \leq a_u M^2 + (1-a_u-g_u)M + g_u$, where $a_u \geq \lambda s_u$ and $g_u \leq ue_0 a_u/\lambda$. Hence, $M_t \leq p(a_u M^2 + (1-a_u-g_u)M + g_u) + (1-p-pe_t)M + pe_t$ $= a_t M^2 + (1-a_t-g_t)M + g_t$, where $a_t = pa_u$ and $g_t = pg_u + pe_t$. Since $p \geq r_t$ and $a_u \geq \lambda s_u \geq \lambda s_{t-1}$, we have $a_t \geq \lambda s_t$. It is easy to check from the definitions that $e_t \leq e_0 s_{t-1}$. Indeed, $\log e_t = -d_0(2N)^t$, while $\log(e_0 s_{t-1}) = \log e_0 + (2N-1)\sum_{j=1}^{t-1} \log e_j$ $= -d_0((2N)^t - 2N + 1)$. Since $g_u \leq ue_0 a_u/\lambda$ and $e_t \leq e_0 s_{t-1} \leq e_0 s_u \leq e_0 a_u/\lambda$, we have $g_t = pg_u + pe_t \leq p(u+1)e_0 a_u/\lambda \leq te_0 a_t/\lambda$.

Therefore, for all t we have $M_t \leq a_t M^2 + (1-a_t-g_t)M + g_t$, where $a_t \geq \lambda s_t$ and $g_t \leq te_0 a_t/\lambda$. Let t be an index with the maximum M_t, i.e., $M = M_t$. Then $M \leq a_t M^2 + (1-a_t-g_t)M + g_t$, hence $a_t M^2 - (a_t+g_t)M + g_t \geq 0$. That is, $(a_t M - g_t)(M-1) \geq 0$. From Claim 3.22, $M < 1$. Therefore, $a_t M \leq g_t$. Thus, $M \leq g_t/a_t \leq te_0/\lambda \leq \varepsilon$. □

This concludes the proof of Theorem 3.19. □

From the constructions in the proof of the theorem we have the following:

Corollary 3.24. *Suppose Algorithm 3.3 outputs the set F when it terminates. Let $S := X - F$.*

1. *There is a randomized static strategy σ for the max player (maximizing non-reachability) such that for all variables $x_i \in S$, we have $(g^*_{\sigma,*})_i > 0$.*

2. *For all* $\varepsilon > 0$*, there is a randomized static strategy* τ_ε*, for the min player (minimizing non-reachability), such that for all variables* $x_i \in F$*,* $(g^*_{*,\tau_\varepsilon})_i \leq \varepsilon$.

Proof. This follows directly from the strategies σ, and τ_ε (for any given $\varepsilon > 0$), constructed in the proof of Theorem 3.19. □

Remark. The static strategies τ_ε from Corollary 3.24(2) can involve probabilities doubly-exponentially small, as a function of the encoding size of the history. However, these probabilities can be encoded in a suitable succinct notation and, hence, τ_ε can be described in a suitable compact form in time polynomial in the encoding size of the input BCSG.

3.6 On the complexity of quantitative problems for BPs

All quantitative decision (e.g., deciding whether the BCSG game value is at least a given probability $p \in (0,1)$) and approximation (i.e., approximating the BCSG game value within a given desired error $\varepsilon > 0$) problems for BCSG reachability games are in PSPACE. This was already known for the reachability objective in the special cases of BPs, BMDPs and BSSGs, and also for the extinction objective in BPs and all their MDPs and (concurrent) game variants.

This upper bound follows, as a corollary from Theorem 3.1, by exploiting the minimax-PPS equations whose greatest (and least) fixed point solution captures the non-reachability (and extinction) values of these games, and by then appealing to PSPACE upper bounds for deciding the Existential Theory of the Reals ([Ren92, Can88], also see Section 2.1 in the Background chapter), in order to decide questions about, and to approximate, the LFP and GFP of such systems of equations. The proof is directly analogous to the proof from [EY08, Theorem 3.3] for the PSPACE upper bound for BCSG extinction games. Given it is an important upper bound for BCSG reachability games, the following is an adapted version of that proof.

Theorem 3.25 (cf. [EY08], Theorem 3.3)**.** *Given a BCSG with minimax-PPS,* $x = P(x)$*, a type* T_v *and a rational probability* $p \in (0,1)$*, there is a* PSPACE *procedure to decide whether* $g^*_v \triangle p$ *(i.e., whether the non-reachability value, starting with an object of type* T_v*, is* $\triangle p$*), where* $\triangle := \{<, \leq, =, \geq, >\}$*. Furthermore, the vector* g^* *of non-reachability values can be approximated to within a given number of bits* j *of precision (*j *is given in unary) in* PSPACE *and in time* $O(j \cdot |P|^{O(n)})$.

Proof. First, it needs to be shown that any of the equations in the minimax-PPS, $x = P(x)$, can be expressed in the ETR. For any $P_i(x)$, $i \in [n]$, of form L or Q, it is clear that the equation $x_i = P_i(x)$ is a multi-variate polynomial equation. As for equations of form M, i.e., $x_i = P_i(x) = Val(A_i(x))$, by the Linear Programming encoding of the minimax theorem for the one-step matrix game $A_i(x)$, the equation can be expressed via the following conjunction of constraints with additional existentially quantified variables $s_i(a_{max})$, $a_{max} \in \Gamma^i_{max}$, and $t_i(a_{min})$, $a_{min} \in \Gamma^i_{min}$, denoting the probabilities for the actions of the two players in the one-step matrix game $A_i(x)$:

$$\forall a_{max} \in \Gamma^i_{max} : \ s_i(a_{max}) \geq 0; \qquad \sum_{a_{max} \in \Gamma^i_{max}} s_i(a_{max}) = 1;$$
$$\forall a_{min} \in \Gamma^i_{min} : \ t_i(a_{min}) \geq 0; \qquad \sum_{a_{min} \in \Gamma^i_{min}} t_i(a_{min}) = 1;$$
$$\forall a_{min} \in \Gamma^i_{min} : \ \sum_{a_{max} \in \Gamma^i_{max}} s_i(a_{max}) \cdot A_i(x)_{(a_{max}, a_{min})} \geq x_i;$$
$$\forall a_{max} \in \Gamma^i_{max} : \ \sum_{a_{min} \in \Gamma^i_{min}} t_i(a_{min}) \cdot A_i(x)_{(a_{max}, a_{min})} \leq x_i$$

Now that the minimax-PPS is expressed in the ETR, one can also encode any question $g^*_v \triangle p$ in the ETR. For instance, $g^*_v \geq p$ can be translated into the ETR formula: $\exists x_1, \dots, x_n \bigwedge_{i \in [n]} (x_i = P_i(x)) \wedge \bigwedge_{i \in [n]} (0 \leq x_i \leq 1) \wedge (x_v \geq p)$. The formula is satisfied if and only if there is a fixed point $g' = P(g')$, $g' \in [0,1]^n$, such that $g'_v \geq p$. But as g^* is the greatest fixed point in $x = P(x)$, then the formula holds if and only if $g^*_v \geq p$.

Another possible decision question can be to determine whether $g^*_v = p$. Consider the ETR formula: $\varphi_1 \equiv \exists x_1, \dots, x_n \bigwedge_{i \in [n]} (x_i = P_i(x)) \wedge \bigwedge_{i \in [n]} (0 \leq x_i \leq 1) \wedge (x_v = p)$. Clearly φ_1 is satisfied if and only if there is a fixed point $g' = P(g')$, $g' \in [0,1]^n$, such that $g'_v = p$. But in order to guarantee that p is the value for the v-th coordinate in the GFP g^*, the following additional ETR formula: $\varphi_2 \equiv \exists x_1, \dots, x_n \bigwedge_{i \in [n]} (x_i = P_i(x)) \wedge \bigwedge_{i \in [n]} (0 \leq x_i \leq 1) \wedge \bigwedge_{i \in [n]} (x_i \geq g'_i) \wedge (x_v > p)$, needs to be checked. φ_2 is false if and only if there is no fixed point $y \in [0,1]^n$ to system $x = P(x)$ such that $y \geq g'$ and $y_v > p$. That is, to decide whether $g^*_v = p$ one needs to make two queries to the ETR decision procedure.

Since the ETR provides a way to decide whether, for some $a, b \in [0,1]$ and for any $v \in [n]$, $a \leq g^*_v \leq b$, a binary search can be performed to achieve approximation of the value g^*_v in the following way. Start with the information that $0 \leq g^*_v \leq 1$. Then at any point if it is known that for some $a, b \in [0,1]$, $a \leq g^*_v \leq b$, perform another ETR query to decide whether $a \le g_v^* \le (a+b)/2$. If yes, then continue the binary search in the interval $[a,(a+b)/2]$, otherwise in the interval $[(a+b)/2,b]$. So if the aim is to compute g_v^* within a precision of j bits, then one needs to perform j number of ETR queries in order to find some $a,b \in [0,1]$ such that $b-a = 1/2^j$ and $g_v^* \in [a,b]$. □

As for lower bounds, it indirectly follows from [EY09, Theorem 5.3] that the quantitative reachability decision problems, even for the special purely probabilistic case of BPs (i.e., no players), are at least as hard as a fundamental problem on arithmetic decision circuits, namely PosSLP. And since the long-standing open problem SQRT-SUM is P-time reducible to the PosSLP problem (see Section 2.1 for descriptions of and relation between the two problems), then quantitative reachability decision problems for BPs are also hard for this major problem in exact numerical computation complexity. This implies that any substantial improvement on PSPACE for such quantitative decision problems and, in fact, even placing these decision problems in the Polynomial Hierarchy would require a major breakthrough on exact numerical computation.

With the purpose of complexity analysis being self-contained here, we provide a proof for this lower bound. This proof is an adaption of [EY09, Theorem 5.3], which showed a reduction from PosSLP to the quantitative *termination* decision problems for 1-exit RMCs (more precisely, to the special case of hierarchical 1-exit RMCs). But it is not immediate to see the consequence for reachability in BPs. In [ESY18], footnote 2 gave a good argument of how the result from [EY09, Theorem 5.3] implies PosSLP-hardness for reachability in BPs. [4]

Theorem 3.26 (cf. [EY09], Theorem 5.3)**.** *The decision problem of determining whether the non-reachability probability is $\ge p$ (or $> p$, etc.) for BPs, for a given rational probability $p \in (0,1)$, is* PosSLP*-hard (and, therefore,* SQRT-SUM*-hard).*

Proof. Let us discuss the reduction from PosSLP. An arithmetic circuit C with inputs $0,1$ and over basis $\{+,*,-\}$ is given. Notice that we can assume w.l.o.g. that there is at most one subtraction gate that can occur as the top gate of the circuit. So the PosSLP problem can be rephrased as the problem of, given two monotone arithmetic circuits S_1, S_2 with inputs $0,1$ and over basis $\{+,*\}$, determine whether $val(S_1) > val(S_2)$,

[4] To summarize the argument from footnote 2 in [ESY18], recall that computing termination probabilities in 1-exit RMCs is equivalent to computing extinction probabilities in BPs, which in turn corresponds to computing the LFP of a PPS, associated with the given BP. So, in the end, [EY09, Theorem 5.3] shows that the problem of, given a start type T_i and a probability $p \in (0,1)$, deciding whether $q_i^* \ge p$ is PosSLP-hard, where q^* is the LFP of the PPS. But since the PPS of the constructed BP in [EY09, Theorem 5.3] has a unique fixed point (i.e., $q^* = g^*$), then the hardness result also applies for the GFP, i.e., the non-reachability probabilities of the BP.

where we denote by $val(X)$ and $val(a)$ the output value of a circuit X and a gate a, respectively. Moreover, w.l.o.g. it can be assumed that the two circuits S_1, S_2 have the same depth, that each level of each of the two circuits consists of either $+$-gates or $*$-gates with inputs from the gates at only the previous level, and that the levels of $+$- and $*$-gates alternate.

Let c be any rational constant in the range $(0,1)$. Analogous to the proof of [EY09, Theorem 5.3], let us construct bottom-up a BP, $\mathcal{A}$, with two types T_i, T_i' for each gate a_i in the circuits S_1 and S_2 such that the non-reachability probabilities g_i and g_i', starting correspondingly at an object of type T_i or T_i', are $\theta_r \cdot val(a_i)$ and $c - g_i$, respectively, where θ_r is a value that depends on the level r of the gate a_i.

First, in the circuits S_1 and S_2, the inputs 0 and 1 can be treated as level-0 gates a_0 and a_{-1}, respectively. Let $\theta_0 = c$. In the BP, $\mathcal{A}$, create the following types and rules (note that, for readers convenience, the right-hand sides of rules are not given as multi-sets), where T_{f^*} is the target type:

$$T_0 \xrightarrow{1} T_{f^*} \qquad T_0' \xrightarrow{c} \emptyset \qquad T_{-1} \xrightarrow{c} \emptyset$$
$$T_{-1}' \xrightarrow{1} T_{f^*} \qquad T_0' \xrightarrow{1-c} T_{f^*} \qquad T_{-1} \xrightarrow{1-c} T_{f^*}$$

Hence, the non-reachability probabilities for these types are: $g_0 = g_{-1}' = 0$ and $g_0' = g_{-1} = c$.

Now consider level $r \geq 1$ of $+$-gates and let $\theta_r = \frac{\theta_{r-1}}{2}$. For any $+$-gate $a_i = a_j + a_k$, create two types T_i, T_i' with the rules:

$$T_i \xrightarrow{1/2} T_j \qquad T_i' \xrightarrow{1/2} T_j'$$
$$T_i \xrightarrow{1/2} T_k \qquad T_i' \xrightarrow{1/2} T_k'$$

Then the non-reachability probability of type T_i is $g_i = \frac{1}{2}(g_j + g_k) = \frac{\theta_{r-1}}{2}(val(a_j) + val(a_k)) = \theta_r val(a_i)$, and the non-reachability probability of type T_i' is $g_i' = \frac{1}{2}(g_j' + g_k') = \frac{1}{2}(c - g_j + c - g_k) = c - g_i$.

Considering level $r \geq 1$ of $*$-gates, let $\rho = \frac{1-c}{2-c^2}$ and $\theta_r = \rho(\theta_{r-1})^2$ and for any gate $a_i = a_j * a_k$, create two types T_i, T_i' with the rules:

$$T_i \xrightarrow{1-\rho} T_{f^*} \qquad T_i' \xrightarrow{1-2\rho} \emptyset \qquad H_j \xrightarrow{(1-c)/2} T_{f^*} \qquad H_k \xrightarrow{(1-c)/2} T_{f^*}$$
$$T_i \xrightarrow{\rho} T_j\, T_k \qquad T_i' \xrightarrow{\rho} T_j'\, H_k \qquad H_j \xrightarrow{c/2} \emptyset \qquad H_k \xrightarrow{c/2} \emptyset$$
$$T_i' \xrightarrow{\rho} T_k'\, H_j \qquad H_j \xrightarrow{1/2} T_j \qquad H_k \xrightarrow{1/2} T_k$$

The non-reachability probability of type T_i is $g_i = \rho g_j g_k = \rho(\theta_{r-1})^2 val(a_j) val(a_k) = \theta_r val(a_i)$. The non-reachability probability of type T_i' is:

$$
\begin{aligned}
g_i' &= 1 - 2\rho + \rho g_j' \left(\frac{c}{2} + \frac{1}{2} g_k\right) + \rho g_k' \left(\frac{c}{2} + \frac{1}{2} g_j\right) \\
&= 1 - 2\rho + \frac{\rho}{2}\Big((c - g_j)(c + g_k) + (c - g_k)(c + g_j)\Big) = 1 - 2\rho + \rho(c^2 - g_j g_k) \\
&= 1 - 2\rho + \rho c^2 - \rho g_j g_k = 1 - \frac{2-2c}{2-c^2} + \frac{(1-c)c^2}{2-c^2} - g_i = c - g_i
\end{aligned}
$$

For the final part of the construction of the BP, $\mathcal{A}$, let a_{m_1} and a_{m_2} be the output gates of S_1 and S_2, respectively, both at the same depth k. Add type T_C to $\mathcal{A}$ with rules: $T_C \xrightarrow{1/2} T_{m_1}$ and $T_C \xrightarrow{1/2} T'_{m_2}$. The non-reachability probability of type T_C is $g_C = \frac{1}{2}(g_{m_1} + g'_{m_2}) = \frac{1}{2}(g_{m_1} + c - g_{m_2}) = \frac{c}{2} + \frac{\theta_k}{2}(val(a_{m_1}) - val(a_{m_2})) = \frac{c}{2} + \frac{\theta_k}{2}(val(S_1) - val(S_2))$. Hence, $g_C > \frac{c}{2}$ if and only if $val(S_1) > val(S_2)$. □

An interesting question is how much the PSPACE upper bounds can be improved for the *approximation* problems. It was shown in [HIJM14] that even for finite-state CSG reachability games, using the standard algorithms for (approximately) solving these games, *value iteration* and *policy iteration*, can be extremely slow in the worst-case: they can take a doubly-exponential number of iterations to obtain any nontrivial approximation, even when the reachability value is 1. Since we know that the problem of computing the reachability values in a BCSG can be rephrased as the problem of computing the greatest fixed point of an associated minimax-PPS, $x = P(x)$, then the above result implies that if we start with the all 1-vector and continuously apply the operator $P(\cdot)$ it can take doubly-exponentially many iterations until the sequence $P^k(\mathbf{1})$, $k \geq 1$ converges within a desired error $\varepsilon > 0$. Furthermore, Frederiksen and Miltersen have shown in [FM13, Theorem 1] that for finite-state CSG reachability games, the game value can be approximated to a desired precision in TFNP[NP]. We do not know an analogous complexity result for quantitative approximation problems for BCSG reachability (or extinction) games, nor do we know PosSLP-hardness (or even SQRT-SUM-hardness) for these approximation problems. These interesting questions are left open.

Finally, the next proposition shows the complexity class FIXP as an upper bound for the problem of computing exact (optimal) reachability probabilities for a BP (and for a minimizing BMDP) (equivalently, computing the greatest fixed point in the associated (max)PPS). It has already been shown in [EY10, Theorem 27] that computing the game values in finite-state CSGs is in FIXP; and in [EY10, Theorem 28] that computing the extinction probabilities in BPs (equivalently, the termination probabilities in 1-exit RMCs, SCFGs and OBPs) is in FIXP.

Proposition 3.27. *The problem of computing the optimal reachability probabilities for a minimizing BMDP, i.e., computing the GFP of a maxPPS, is in* FIXP.

Proof. Recall that in order for a search problem to be shown that it belongs in the complexity class FIXP, it needs to be expressible as a fixed point problem for a continuous function over algebraic circuits with basis $\{+,-,*,/,\min,\max,\sqrt[k]{}\}$ and with rational constants, such that the set of solutions to the given problem is precisely the set of fixed points of the function.

By Theorem 3.1, given a minimizing BMDP with the reachability objective, one can construct a corresponding maxPPS, $x = F(x)$, such that the greatest fixed point captures the vector of optimal non-reachability probabilities. $F(\cdot)$ is a monotone function over the unit n-cube, satisfying the FIXP class requirements for the function and domain. However, the issue here is that there may be multiple fixed points, but only the GFP is a solution to our problem. So if one can show that this system, $x = F(x)$, can be modified in such a way that the GFP is the unique fixed point, then the inclusion in the complexity class FIXP follows immediately.

Let us remove all variables x_i such that the optimal non-reachability probability, starting at an object of corresponding type T_i, is 1. This is done in P-time using Algorithm 3.1. Let us denote by $x = P(x)$ the reduced maxPPS system on the remaining variables. Note that the GFP, g^*, of $x = P(x)$ satisfies $g^* < \mathbf{1}$. By Lemma 3.9 we know that, since GFP $g^* < \mathbf{1}$, then g^* is in fact the unique fixed point of $x = P(x)$ in $[0,1]^n$. That concludes the proof. □

As pointed in Section 2.1, computing (respectively, approximating) Nash Equilibrium for 3 or more players is FIXP-complete (respectively, FIXP_a-complete). Therefore, the decision and approximation questions for the reachability objective for minimizing BMDPs (equivalently, the decision and approximation questions for the GFP of a maxPPS) reduce to the decision and approximation questions for the Nash Equilibria problem for 3 or more players. However, for the approximation questions, there is already a P-time procedure to approximate (within a given desired error $\varepsilon > 0$) the GFP of a maxPPS and provide a deterministic static ε-optimal strategy (see [ESY18, Theorem 6.3]). Same holds for minPPSs, i.e., there is a P-time procedure to approximate (within a given desired error $\varepsilon > 0$) the GFP of a minPPS (equivalently, approximate the optimal reachability probabilities for a maximizing BMDP) and provide a randomized static ε-optimal stategy (see [ESY18, Theorems 7.1, 8.8]). It is an open question whether approximating the GFP of a minimax-PPS (equivalently, approximating the reachability values for a BCSG) is in FIXP_a. On the other hand, it has been shown in an unpublished manuscript ([ESY14]) that approximating the LFP of a minimax-PPS (equivalently, approximating the extinction values for a BCSG or the termination values for an 1-exit RCSG, a CF-CSG and an OBCSG) is in FIXP_a.

Chapter 4

Multi-Objective Reachability for Ordered Branching MDPs

In this chapter we focus on *multi-objective reachability* questions in the context of the Ordered Branching MDP model.

The single-target reachability objective for OBMDPs amounts to optimizing (maximizing or minimizing) the probability that, starting at a given starting (root) non-terminal, the generated tree contains some given target non-terminal. As mentioned in the related work (see subsection 2.6.1), this objective has already been thoroughly studied for BMDPs, as well as for BPs and for the (concurrent) stochastic game generalizations of BMDPs (Chapter 3 and [ESY18]). Moreover, as it turned out (in Proposition 2.4), there is really no difference between BMDPs and OBMDPs when it comes to the single-target reachability objective: all the algorithmic results from [ESY18] and Chapter 3 ([EMSY19]) carry over, *mutatis mutandis*, for OBMDPs, and for their purely probabilistic OBP version and stochastic game generalizations.

A natural generalization of single-target reachability is multi-objective reachability, where the goal is to optimize each of the respective probabilities that the generated tree satisfies each of several given generalized reachability objectives over different target non-terminals. Of course, there may be trade-offs between these different objectives.

Our main concern in this chapter is the specific *qualitative* multi-objective reachability questions, where the aim is to determine whether there is a strategy that guarantees that each of a given set of target non-terminals is almost-surely (respectively, limit-surely) contained in the generated tree, i.e., with probability 1 (respectively, with probability arbitrarily close to 1). In fact, we show that in this context the *almost-sure* and *limit-sure* problems do not coincide. That is, there are OBMDPs for which there is no single strategy that achieves probability exactly 1 for reaching all targets, but where nevertheless, for every $\varepsilon > 0$, there is a strategy that guarantees a probability $\geq 1-\varepsilon$ of reaching all targets.

By contrast, for both BMDPs and OBMDPs, for single-target reachability, the *qualitative* almost-sure and limit-sure problems do coincide: there is a strategy that guarantees reaching the target non-terminal with probability 1 if and only if there is a sequence of strategies that guarantees reaching the target with probabilities arbitrarily close to 1 ([ESY18]).

We give two separate algorithms for almost-sure and limit-sure multi-target reachability. For the *almost-sure* problem, we are given an OBMDP, a start non-terminal, and a set of target non-terminals, and we must decide whether there exists a strategy using which the process generates, with probability 1, a tree that contains all the given target non-terminals. If the answer is "yes", the algorithm can also be easily augmented to construct (proof shows how) a randomized witness strategy that achieves this.[1] The algorithm for the *limit-sure* problem decides whether the supremum probability of generating a tree that contains all the given target non-terminals is 1. If the answer is "yes", the algorithm can also be easily augmented to construct (proof shows how), given any $\varepsilon > 0$, a randomized non-static strategy that guarantees probability $\geq 1-\varepsilon$.

Both algorithms run in time $2^{O(k)} \cdot |\mathcal{A}|^{O(1)}$, where $|\mathcal{A}|$ is the total bit encoding length of the given OBMDP, $\mathcal{A}$, and $k = |K|$ is the size of the given set K of target non-terminals. Hence, they run in polynomial time when k is fixed and also are fixed-parameter tractable (FPT) with respect to k. Moreover, we show that the qualitative almost-sure (and limit-sure) multi-target reachability decision problem is in general NP-hard, when k is not fixed.

Going beyond the goal of assuring probability 1 of reaching each of a set of target non-terminals, we also consider more general qualitative multi-objective reachability problems, where we are given a set of target non-terminals, K, and where, for each target non-terminal T_q ($q \in K$), we are also given a 0/1 probability $b_q \in \{0,1\}$, and an inequality $\Delta_q \in \{<,=,>\}$, and where we wish to decide whether the controller has a single strategy using which, for all $q \in K$ the probability that the generated tree contains the non-terminal T_q is $\Delta_q b_q$. We show that in some special cases these problems are efficiently decidable. However, we leave open the decidability of the most

[1] This strategy is, however, necessarily not "static", meaning it must actually use the ancestor history: the action distribution cannot be defined solely based on which non-terminal is being expanded.

general case of *arbitrary* Boolean combinations of such qualitative reachability and non-reachability queries over different target non-terminals. Furthermore, we leave open all (both decision and approximation) *quantitative* multi-objective reachability questions, including when the goal is to approximate the tradeoff *pareto curve* of optimal probabilities for different reachability objectives. These are intriguing questions for future research and we come back to them in Chapter 5.

Before we move on with the chapter, we would like to briefly and informally recap the differences between BMDPs and OBMDPs. Although both models are similar, the seemingly small differences between them are crucial. In BMDPs, the children that each rule generates is a multi-set over the types and there is no ordering among the children. However, in OBMDPs, there is an ordering among the non-terminals generated by a rule and this turns out to be beneficial.

As already pointed in subsection 2.4.1 computing the optimal (single-target) reachability probabilities in OBMDPs is equivalent to computing the optimal (single-target) reachability probabilities in BMDPs. And the same holds for the objective of extinction/termination. And here is where the key differences between these two models manifests. In the context of BMDPs, we need the more general notion of a strategy (i.e., strategy having the information of the entire finite tree up to the current generation) in order to obtain even the qualitative almost-sure winning strategies, due to the lack of ordering among objects in a generation. However, in the context of OBMDPs, such strategies can be implemented even with the restricted notion of ancestral strategies, due to the fact that there is indeed an ordering among the non-terminals.

There is no “suitable” definition of a strategy or a history for the models of branching processes (also discussed in Section 5.1). But it is quite interesting that we show that, when ordering in the tree is introduced, the more general notion of a history is not more powerful than an ancestor history for the objectives of single-target reachability and termination. The latter definition of a history may reveal less information, but at the same time it is less computationally expensive to implement.

This motivated us to introduce the OBMDP model. This model carries with it the idea that, for each object of the current generation, it is irrelevant for the player to have information regarding what is happening in other parts of the tree. And there may be other objectives, where such a property facilitates the analysis.

Organization of the chapter. Section 4.1 shows NP-hardness for qualitative multi-target reachability decision problems. Section 4.2 gives an algorithm for determining the non-terminals starting from which, regardless of the strategy, there is a zero probability that all of the given target non-terminals are in the generated tree. Sections 4.3 and 4.4 provide, respectively, the algorithms for the limit-sure and almost-sure multi-target reachability problems. Section 4.5 considers other certain cases of qualitative multi-objective reachability.

4.1 On the complexity of multi-target reachability for OBMDPs

Before we continue with the algorithmic results, let us observe that the qualitative (both almost-sure and limit-sure) multi-target reachability problems are in general NP-hard (coNP-hard), if the size of the set K of target non-terminals is not bounded by a fixed constant.

Proposition 4.1.

(1.) The following two problems are both NP-hard: given an OBMDP, a set $K \subseteq [n]$ of target non-terminals and a starting non-terminal $T_i \in V$, decide whether: (i) $\exists \sigma \in \Psi : Pr^{\sigma}_{T_i}[\bigcap_{q \in K} Reach(T_q)] = 1$, and (ii) $Pr^{*}_{T_i}[\bigcap_{q \in K} Reach(T_q)] = 1$.

(2.) The following problem is coNP-hard: given an OBP (i.e., an OBMDP with no controlled non-terminals, and hence with only one trivial strategy σ), a set $K \subseteq [n]$ of target non-terminals and a starting non-terminal $T_i \in V$, decide whether $Pr^{\sigma}_{T_i}[\bigcap_{q \in K} Reach(T_q)] = 0$.

Proof. For (1.) we reduce from 3-SAT, and for (2.) from the complement problem (i.e., deciding unsatisfiability of a 3-CNF formula). The reductions are nearly identical, so we describe them both together. Consider a 3-CNF formula over variables $\{x_1, \dots, x_n\}$:

$$\bigwedge_{q \in [m]} (l_{q,1} \vee l_{q,2} \vee l_{q,3})$$

where every $l_{q,j}$ is either x_r or $\neg x_r$ for some $r \in [n]$. We construct an OBMDP as follows: to each clause $q \in [m]$ we associate a target non-terminal R_q with a single associated rule $R_q \xrightarrow{1} \varnothing$; for each variable $x_r, r \in [n]$, we associate two purely probabilistic non-terminals T_{r_a}, T_{r_b}, and

- for (1.), a controlled non-terminal C_r with rules $C_r \xrightarrow{a} T_{r_a}$ and $C_r \xrightarrow{b} T_{r_b}$, or

- for (2.), a probabilistic non-terminal C_r with rules $C_r \xrightarrow{1/2} T_{r_a}$ and $C_r \xrightarrow{1/2} T_{r_b}$.

For each non-terminal T_{r_a}, $r \in [n]$, we would in principle like to create a single rule, with probability 1, whose right-hand side consists of the following non-terminals (in any order): $\{R_q \mid \exists j \in \{1,2,3\}\ s.t.\ l_{q,j} = x_r\}$, as well as the non-terminal C_{r+1} if $r < n$; likewise, for each non-terminal T_{r_b}, $r \in [n]$, we would like to create a single rule, with probability 1, whose right-hand side consists of $\{R_q \mid \exists j \in \{1,2,3\}\ s.t.\ l_{q,j} = \neg x_r\}$, as well as C_{r+1} if $r < n$.

However, due to the simple normal form we have adopted in our definition of OBMDPs, such rules need to be "expanded" (as shown in Proposition 2.3) into a sequence of rules whose right-hand side has length ≤ 2, using auxiliary non-terminals. So, for example, instead of a single rule of the form $T_{1_b} \xrightarrow{1} R_2 R_3 R_4 C_2$, we will have the following rules (using auxiliary non-terminals $T^j_{1_b}$): $T_{1_b} \xrightarrow{1} R_2\, T^1_{1_b}$, $T^1_{1_b} \xrightarrow{1} R_3\, T^2_{1_b}$, and $T^2_{1_b} \xrightarrow{1} R_4\, C_2$. See Figure 4.1 for an example.

$C_1 \xrightarrow{a} T_{1_a}$	$T_{1_a} \xrightarrow{1} R_1\, C_2$	$C_2 \xrightarrow{a} T_{2_a}$	$T_{2_a} \xrightarrow{1} R_2\, T^1_{2_a}$	$C_3 \xrightarrow{a} T_{3_a}$	$T_{3_a} \xrightarrow{1} R_1\, R_3$
$C_1 \xrightarrow{b} T_{1_b}$	$T_{1_b} \xrightarrow{1} R_2\, T^1_{1_b}$	$C_2 \xrightarrow{b} T_{2_b}$	$T^1_{2_a} \xrightarrow{1} R_3\, C_3$	$C_3 \xrightarrow{b} T_{3_b}$	$T_{3_b} \xrightarrow{1} R_2\, R_4$
	$T^1_{1_b} \xrightarrow{1} R_3\, T^2_{1_b}$		$T_{2_b} \xrightarrow{1} R_1\, T^1_{2_b}$		
	$T^2_{1_b} \xrightarrow{1} R_4\, C_2$		$T^1_{2_b} \xrightarrow{1} R_4\, C_3$		

Figure 4.1: Reduction example: an OBMDP obtained from the 3-SAT formula $(x_1 \vee \neg x_2 \vee x_3) \wedge (\neg x_1 \vee x_2 \vee \neg x_3) \wedge (\neg x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee \neg x_2 \vee \neg x_3)$. This construction is for the problems in (1.); the construction for the problem in (2.) is very similar, with the controlled non-terminals $C_r, r \in [n]$ changed to purely probabilistic non-terminals instead (with $1/2$ probability on each of their two rules).

This reduction closely resembles a well-known reduction ([SC85, Theorem 3.5]) for NP-hardness of model checking eventuality formulas in linear temporal logic. The immediate children of the branching non-terminals T_{r_a} and T_{r_b} keep track of which clauses are satisfied under each of the two truth assignments to the variable x_r ('true' corresponds to T_{r_a}, and 'false' corresponds to T_{r_b}). In fact, for the OBMDP obtained for the problems in (1.), there is a one-to-one correspondence between truth assignments to all variables of the formula and deterministic static strategies.

It follows that, for the OBMDP in statement (1.), if there exists a satisfying truth assignment for the formula, then starting at a non-terminal C_1, there exists a (deterministic and static) strategy σ' for the player such that $Pr^{\sigma'}_{C_1}[\bigcap_{q\in[m]} Reach(R_q)] = 1$.

Otherwise, if the formula is unsatisfiable, then the claim is that for every $\sigma \in \Psi$: $Pr^{\sigma}_{C_1}[\bigcap_{q\in[m]} Reach(R_q)] = 0$. (And hence, that $Pr^{*}_{C_1}[\bigcap_{q\in[m]} Reach(R_q)] = 0 < 1$.) To see this, note that an arbitrary (possibly randomized, and not necessarily static) strategy in the constructed OBMDP corresponds to a (possibly correlated) probability distribution on assignments of truth values to the variables in the corresponding formula. (The distribution may be correlated, because the strategy may be non-static, but this doesn't matter.) So if the formula is unsatisfiable, then under any strategy for the player (i.e., any probability distribution on assignments of truth values), there is probability 0 that the generated play (tree) contains all target non-terminals (respectively, that the random truth assignment satisfies all clauses in the formula).

For the problem in (2.), it follows from the same arguments that the formula is unsatisfiable if and only if $Pr^{\sigma}_{C_1}[\bigcap_{q\in[m]} Reach(R_q)] = 0$ (where σ is just the trivial strategy, since there are no controlled non-terminals in the OBP obtained for (2.)). □

Throughout the next three sections we will provide algorithms for the problems of Proposition 4.1. As a consequence from the running time of these algorithms, it follows that there is an EXPTIME upper bound on the problems. We leave open the question of whether this upper bound can be improved.

Before we continue with these algorithms, we provide some more notation in the context of OBMDPs, needed for this chapter. We shall hereafter use the notation $T_i \rightarrow T_j$ (respectively, $T_i \not\rightarrow T_j$), to denote that for non-terminal T_i there *exists* (respectively, there does *not* exist) either an associated (controlled) rule $T_i \xrightarrow{a} T_j$, where $a \in \Gamma^i$, or an associated probabilistic rule $T_i \xrightarrow{p_{i,j}} T_j$ with a positive probability $p_{i,j} > 0$. Similarly, let $T_i \rightarrow \varnothing$ (respectively, $T_i \not\rightarrow \varnothing$) denote that the rule $T_i \xrightarrow{p_{i,0}} \varnothing$ has a positive probability $p_{i,0} > 0$ (respectively, has a probability $p_{i,0} = 0$).

Definition 12. *The **dependency graph** of a SNF-form OBMDP, $\mathcal{A}$, is a directed graph that has a node T_i for each non-terminal T_i, and contains an edge (T_i, T_j) if and only if: either $T_i \rightarrow T_j$ or there is a rule $T_i \xrightarrow{1} T_j\ T_r$ or a rule $T_i \xrightarrow{1} T_r\ T_j$ in $\mathcal{A}$.*

Throughout this paper, for (SNF-form) OBMDP, $\mathcal{A}$, with non-terminals set V, we let $G = (U, E)$, with $U = V$, denote the dependency graph of $\mathcal{A}$ and let $G[C]$ denote the subgraph of G induced by the subset $C \subseteq U$ of nodes (non-terminals).

Sometimes when the specific OBMDP, $\mathcal{A}$, is not clear from the context, we use $\mathcal{A}$ as a superscript to specify the OBMDP in our notations. So, for instance, $\Psi^{\mathcal{A}}$ is the set of all strategies for $\mathcal{A}$; $G^{\mathcal{A}}$ is the dependency graph of $\mathcal{A}$; and $Pr_{T_i}^{\sigma,\mathcal{A}}[\mathcal{F}]$ is the probability of event $\mathcal{F}$, starting at a non-terminal T_i, under strategy σ, in $\mathcal{A}$.

We also extend the notation regarding probabilities of properties to "start" at a given ancestor history. That is, for an ancestor history h, we use $Pr_h^{\sigma,\mathcal{A}}[\mathcal{F}]$ to denote the conditional probability that, using $\sigma \in \Psi^{\mathcal{A}}$, conditioned on the event that there is a node in the play whose ancestor history is h, the *subplay* rooted at $\texttt{current}(h)$, is in the set $\mathcal{F}$. Whenever we use the notation $Pr_h^{\sigma,\mathcal{A}}[\mathcal{F}]$, the underlying conditional probability will be well defined. Again, the superscript $\mathcal{A}$ will be omitted when clear from context.

Note that one ancestor history h can be a prefix of another ancestor history. We use the notation $h' := h(x,T_i)$, for some $x \in \{l,r,u\}$, to denote that h is the immediately prior ancestor history to h', which is obtained by concatenating the pair (x,T_i) at the end of h.

Definition 13. *For a directed graph $G = (U,E)$, and a partition of its vertices $U = (U_1,U_P)$, an **end-component** is a set of vertices $C \subseteq U$ such that $G[C]$: (1) is strongly connected; (2) for all $u \in U_P \cap C$ and all $(u,u') \in E$, $u' \in C$; (3) and if $C = \{u\}$ (i.e., $|C| = 1$), then $(u,u) \in E$. A **maximal end-component (MEC)** is an end-component not contained in any larger end-component. A **MEC-decomposition** is a partition of the graph into MECs and nodes that do not belong to any MEC.*

MECs are disjoint and the unique MEC-decomposition of such a directed graph G (with partitioned nodes) can be computed in P-time ([CY98]).[2] More recent work provides more efficient algorithms for MEC-decomposition (see [CH14]). We will also be using the notion of a strongly connected component (SCC), which can be defined as a MEC where condition (2) from Definition 13 above is not required. It is also well-known that a SCC-decomposition of a directed graph can be done in linear time.

For our setting here, given a SNF-form OBMDP with its dependency graph $G = (U,E)$, $U = V$, the partition of U that we will use is the following: $U_P := \{T_i \in U \mid T_i$ is of L-form$\}$ and $U_1 := \{T_i \in U \mid T_i$ is of M-form or Q-form$\}$.

Before we move on with the algorithmic sections let us provide an OBMDP example where almost-sure multi-target reachability is satisfied. Example 2.1 demonstrated an OBMDP example where almost-sure multi-target reachability is not satisfied, but limit-sure multi-target reachability is satisfied. Both of these examples give a rough idea of the properties non-terminals have and the type of strategies the player utilizes for almost-sure or limit-sure multi-target reachability in OBMDPs. The proofs of the

[2] In [CY98], maximal end-components are referred to as *closed components*.

algorithms in the next sections will provide a clear picture of how to compute the winning non-terminals and how to construct the necessary strategies.

Example 4.1. *OBMDP example demonstrating almost-sure multi-target reachability.*

$$C \xrightarrow{c} M \qquad T \xrightarrow{1} C\,M \qquad M \xrightarrow{a} A \qquad A \xrightarrow{1/2} R_1 \qquad B \xrightarrow{1/2} R_2$$

$$C \xrightarrow{d} T \qquad T' \xrightarrow{1} R_1\,R_2 \qquad M \xrightarrow{b} B \qquad A \xrightarrow{1/2} \varnothing \qquad B \xrightarrow{1/2} \varnothing$$

Consider the OBMDP above with non-terminals set $\{C, T, T', M, A, B, R_1, R_2\}$, where R_1 and R_2 are the target non-terminals and C and M are the "controlled" non-terminals. Clearly, starting at a non-terminal T', both targets are immediately reached in the next step. There is a strategy σ for the player such that, starting at a non-terminal C, it follows that $Pr^{\sigma}_{C}[Reach(R_1) \cap Reach(R_2)] = 1$. The same strategy σ also satisfies almost-sure multi-target reachability for non-terminal T.

To see this, consider the following strategy σ: in every copy of non-terminal C, let σ choose deterministically action d; and in every copy of non-terminal M, let σ choose uniformly at random between actions a and b. Note that starting at a non-terminal C, under σ, infinitely often a copy of non-terminal T is generated and each such copy generates an independent copy of non-terminal M, which has a positive probability bounded away from zero to reach any of the two targets. Hence, with probability 1 both target non-terminals are reached. □

4.2 Algorithm for deciding $\max_{\sigma} Pr^{\sigma}_{T_i}[\bigcap_{q \in K} Reach(T_q)] \stackrel{?}{=} 0$

In this section we present an algorithm that, given an OBMDP and a set $K \subseteq [n]$ of $k = |K|$ target non-terminals, computes, for every subset of target non-terminals $K' \subseteq K$, the set $Z_{K'} \subseteq V$ of non-terminals such that, starting at a non-terminal $T_i \in Z_{K'}$, using any strategy σ, the probability that the generated play contains a copy of every non-terminal in set K' is 0. In other words, Algorithm 4.1 computes, $\forall K' \subseteq K$, the set $Z_{K'} := \{T_i \in V \mid \forall \sigma \in \Psi : \; Pr^{\sigma}_{T_i}[\bigcap_{q \in K'} Reach(T_q)] = 0\}$. The algorithm uses as a preprocessing step an algorithm from [ESY18, Proposition 4.1], which is a special case version of Algorithm 3.1. Namely, let us denote by W_q the set $\{T_q\} \cup \{T_i \in V \mid \exists \sigma \in \Psi : Pr^{\sigma}_{T_i}[Reach(T_q)] > 0\}$. We can compute, for each $q \in K$, the set W_q in P-time using the algorithm from [ESY18, Proposition 4.1], together with a single deterministic static witness strategy for every non-terminal in set W_q. Let K'_{-i} denote the set $K' - \{i\}$.

Algorithm 4.1 Algorithm for computing the set $\{T_i \in V \mid \forall\sigma \in \Psi : Pr^\sigma_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = 0\}$ for every subset of target non-terminals $K' \subseteq K$ in a given OBMDP.

I. Initialize $\bar{Z}_{\{q\}} := W_q$ and $Z_{\{q\}} := V - W_q$, for each $q \in K$. Let $\bar{Z}_\emptyset := V$ and $Z_\emptyset := \emptyset$.

II. For $l = 2 \ldots k$:

For every subset of target non-terminals $K' \subseteq K$ of size $|K'| = l$:

1. Initialize $\bar{Z}_{K'} := \{T_i \in V \mid$ one of the following holds:
   - T_i is of L-form where $i \in K'$ and $T_i \to T_j$, $T_j \in \bar{Z}_{K'_{-i}}$.
   - T_i is of M-form where $i \in K'$ and $\exists a' \in \Gamma^i : T_i \xrightarrow{a'} T_j$, $T_j \in \bar{Z}_{K'_{-i}}$.
   - T_i is of Q-form ($T_i \xrightarrow{1} T_j\, T_r$) where $i \in K'$ and $\exists K_L \subseteq K'_{-i} : T_j \in \bar{Z}_{K_L} \wedge T_r \in \bar{Z}_{K'_{-i} - K_L}$.
   - T_i is of Q-form ($T_i \xrightarrow{1} T_j\, T_r$) and $\exists K_L \subset K'$ ($K_L \neq \emptyset$) $: T_j \in \bar{Z}_{K_L} \wedge T_r \in \bar{Z}_{K' - K_L}$.$\}$
2. Repeat until no change has occurred to $\bar{Z}_{K'}$:
   (a) add $T_i \notin \bar{Z}_{K'}$ to $\bar{Z}_{K'}$, if of L-form and $T_i \to T_j$, $T_j \in \bar{Z}_{K'}$.
   (b) add $T_i \notin \bar{Z}_{K'}$ to $\bar{Z}_{K'}$, if of M-form and $\exists a' \in \Gamma^i : T_i \xrightarrow{a'} T_j$, $T_j \in \bar{Z}_{K'}$.
   (c) add $T_i \notin \bar{Z}_{K'}$ to $\bar{Z}_{K'}$, if of Q-form ($T_i \xrightarrow{1} T_j\, T_r$) and $T_j \in \bar{Z}_{K'} \vee T_r \in \bar{Z}_{K'}$.
3. $Z_{K'} := V - \bar{Z}_{K'}$.

Proposition 4.2. *Algorithm 4.1 computes, given an OBMDP, $\mathcal{A}$, and a set $K \subseteq [n]$ of $k = |K|$ target non-terminals, for every subset of target non-terminals $K' \subseteq K$, the set $Z_{K'} := \{T_i \in V \mid \forall\sigma \in \Psi : Pr^\sigma_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = 0\}$. The algorithm runs in time $4^k \cdot |\mathcal{A}|^{O(1)}$. The algorithm can also be augmented to compute a deterministic (non-static) strategy $\sigma'_{K'}$ and a rational value $b_{K'} > 0$, such that for all $T_i \notin Z_{K'}$, $Pr^{\sigma'_{K'}}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] \geq b_{K'} > 0$.*

Proof. The running time of the algorithm follows from the facts that step II. executes for 2^k iterations and inside each iteration, step II.1. requires time at most $2^k \cdot |\mathcal{A}|^{O(1)}$ and the loop at step II.2. executes in time at most $|\mathcal{A}|^{O(1)}$.

We need to prove that for every $K' \subseteq K : T_i \in Z_{K'}$ if and only if $\forall\sigma \in \Psi$: $Pr^\sigma_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = 0 \Leftrightarrow Pr^\sigma_{T_i}[\bigcup_{q\in K'} Reach^\complement(T_q)] = 1$ (or equivalently, that $T_i \in \bar{Z}_{K'}$ if and only if $\exists\sigma'_{K'} \in \Psi : Pr^{\sigma'_{K'}}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] > 0$). We in fact show that there is a value $b_{K'} > 0$ and a strategy $\sigma'_{K'} \in \Psi$ such that $T_i \in \bar{Z}_{K'}$ if and only if

$Pr_{T_i}^{\sigma'_{K'}}[\bigcap_{q\in K'} Reach(T_q)] \geq b_{K'}$. We analyse this by a double induction with the top-layer induction based on the size of set K', or in other words the time of constructing set $\bar{Z}_{K'}$. Clearly for the base case (step I.) of a single target non-terminal $T_q, q \in K$, by the P-time algorithm from [ESY18, Proposition 4.1], there is a (deterministic static) strategy $\sigma'_{\{q\}}$ for the player and a value $b_{\{q\}} > 0$ where $T_i \in \bar{Z}_{\{q\}}$ if and only if $Pr_{T_i}^{\sigma'_{\{q\}}}[Reach^{\complement}(T_q)] \leq 1 - b_{\{q\}} < 1 \Leftrightarrow Pr_{T_i}^{\sigma'_{\{q\}}}[Reach(T_q)] \geq b_{\{q\}} > 0$. Now, constructing set $\bar{Z}_{K'}$ for a subset $K' \subseteq K$ of target non-terminals of size l, assume that for each $K'' \subset K'$ of size $\leq l-1$, there is a strategy $\sigma'_{K''}$ for the player and a value $b_{K''} > 0$ such that for all $T_j \in \bar{Z}_{K''}$, $Pr_{T_j}^{\sigma'_{K''}}[\bigcup_{q\in K''} Reach^{\complement}(T_q)] \leq 1 - b_{K''} < 1 \Leftrightarrow Pr_{T_j}^{\sigma'_{K''}}[\bigcap_{q\in K''} Reach(T_q)] \geq b_{K''} > 0$. And for all $T_j \in Z_{K''}$, it holds that $\forall \sigma \in \Psi: Pr_{T_j}^{\sigma}[\bigcap_{q\in K''} Reach(T_q)] = 0$.

First, let us prove the direction where there exists $\sigma'_{K'} \in \Psi$ such that, if $T_i \in \bar{Z}_{K'}$, then $Pr_{T_i}^{\sigma'_{K'}}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \leq 1 - b_{K'} < 1 \Leftrightarrow Pr_{T_i}^{\sigma'_{K'}}[\bigcap_{q\in K'} Reach(T_q)] \geq b_{K'} > 0$, for some value $b_{K'} > 0$. We use a second (nested) induction, based on the iteration in which non-terminal T_i was added to set $\bar{Z}_{K'}$. Consider the base case where T_i is a non-terminal added to set $\bar{Z}_{K'}$ at the initialization step II.1.

(i) Suppose T_i is of L-form where $i \in K'$ (i.e., T_i is a target non-terminal in set K') and $T_i \to T_j$, $T_j \in \bar{Z}_{K'_{-i}}$, where by induction $\exists \sigma'_{K'_{-i}} \in \Psi: Pr_{T_j}^{\sigma'_{K'_{-i}}}[\bigcap_{q\in K'_{-i}} Reach(T_q)] \geq b_{K'_{-i}}$, for some value $b_{K'_{-i}} > 0$. Due to the fact that the play up to (and including) a copy of non-terminal $T_i, i \in K'$ has already reached the target T_i and using strategy $\sigma'_{K'_{-i}}$ from the next generation as if the play starts in it, it follows that there exists a strategy $\sigma'_{K'}$ such that, for an ancestor history $h := T_i(u, T_j)$:

$$Pr_{T_i}^{\sigma'_{K'}}\Big[\bigcap_{q\in K'} Reach(T_q)\Big] = Pr_{T_i}^{\sigma'_{K'}}\Big[\bigcap_{q\in K'_{-i}} Reach(T_q) \;\Big|\; Reach(T_i)\Big] \cdot Pr_{T_i}^{\sigma'_{K'}}\Big[Reach(T_i)\Big]$$
$$= Pr_{T_i}^{\sigma'_{K'}}\Big[\bigcap_{q\in K'_{-i}} Reach(T_q)\Big] \geq p_{i,j} \cdot Pr_{h}^{\sigma'_{K'}}\Big[\bigcap_{q\in K'_{-i}} Reach(T_q)\Big]$$
$$= p_{i,j} \cdot Pr_{T_j}^{\sigma'_{K'_{-i}}}\Big[\bigcap_{q\in K'_{-i}} Reach(T_q)\Big] \geq p_{i,j} \cdot b_{K'_{-i}} > 0$$

where $p_{i,j} > 0$ is the probability of the rule $T_i \xrightarrow{p_{i,j}} T_j$. So let $b^i_{K'} := p_{i,j} \cdot b_{K'_{-i}}$.

(ii) Suppose T_i is of M-form where $i \in K'$ and $\exists a' \in \Gamma^i: T_i \xrightarrow{a'} T_j$, $T_j \in \bar{Z}_{K'_{-i}}$. Again let $h := T_i(u, T_j)$. By combining the witness strategy $\sigma'_{K'_{-i}}$ from the induction assumption for a starting non-terminal T_j with the initial local choice of choosing deterministically action a' starting at a non-terminal T_i, we obtain a combined strategy $\sigma'_{K'}$, such that starting at a (target) non-terminal T_i, we satisfy

$Pr^{\sigma'_{K'}}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = Pr^{\sigma'_{K'}}_{h}[\bigcap_{q\in K'_{-i}} Reach(T_q)] = Pr^{\sigma'_{K'_{-i}}}_{T_j}[\bigcap_{q\in K'_{-i}} Reach(T_q)] \geq b_{K'_{-i}} > 0$. So let $b^i_{K'} := b_{K'_{-i}}$.

(iii) Suppose T_i is of Q-form (i.e., $T_i \stackrel{1}{\to} T_j\ T_r$) and there exists a proper split of the target non-terminals from set K', implied by $K_L \subset K'$ (where $K_L \neq \emptyset$) and $K' - K_L$, such that $T_j \in \bar{Z}_{K_L} \wedge T_r \in \bar{Z}_{K'-K_L}$. So, by the inductive assumption, for some values $b_{K_L}, b_{K'-K_L} > 0$, $\exists \sigma'_{K_L} \in \Psi :\ Pr^{\sigma'_{K_L}}_{T_j}[\bigcap_{q\in K_L} Reach(T_q)] \geq b_{K_L} > 0$ and $\exists \sigma'_{K'-K_L} \in \Psi :\ Pr^{\sigma'_{K'-K_L}}_{T_r}[\bigcap_{q\in K'-K_L} Reach(T_q)] \geq b_{K'-K_L} > 0$. Let $h_l := T_i(l, T_j)$ and $h_r := T_i(r, T_r)$. Hence, by combining the two strategies σ'_{K_L} and $\sigma'_{K'-K_L}$ to be used from the next generation from the left and right child, respectively, as if the play starts in them, it follows that $\exists \sigma'_{K'} \in \Psi : Pr^{\sigma'_{K'}}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] \geq Pr^{\sigma'_{K'}}_{h_l}[\bigcap_{q\in K_L} Reach(T_q)] \cdot Pr^{\sigma'_{K'}}_{h_r}[\bigcap_{q\in K'-K_L} Reach(T_q)] = Pr^{\sigma'_{K_L}}_{T_j}[\bigcap_{q\in K_L} Reach(T_q)] \cdot Pr^{\sigma'_{K'-K_L}}_{T_r}[\bigcap_{q\in K'-K_L} Reach(T_q)] \geq b_{K_L} \cdot b_{K'-K_L} > 0$, and so let $b^i_{K'} := b_{K_L} \cdot b_{K'-K_L}$.

(iv) Suppose T_i is of Q-form (i.e., $T_i \stackrel{1}{\to} T_j\ T_r$) where $i \in K'$ and there exists a split of the target non-terminals from set K'_{-i}, implied by $K_L \subseteq K'_{-i}$ and $K'_{-i} - K_L$, such that $T_j \in \bar{Z}_{K_L} \wedge T_r \in \bar{Z}_{K'_{-i}-K_L}$. Combining in the same way as in (iii) above the two witness strategies σ'_{K_L} and $\sigma'_{K'_{-i}-K_L}$ from the induction assumption for non-terminals T_j and T_r, and the fact that the play starts in the target non-terminal T_i ($i \in K'$), it follows that there exists $\sigma'_{K'} \in \Psi$ such that $Pr^{\sigma'_{K'}}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = Pr^{\sigma'_{K'}}_{T_i}[\bigcap_{q\in K'_{-i}} Reach(T_q)] \geq Pr^{\sigma'_{K_L}}_{T_j}[\bigcap_{q\in K_L} Reach(T_q)] \cdot Pr^{\sigma'_{K'_{-i}-K_L}}_{T_r}[\bigcap_{q\in K'_{-i}-K_L} Reach(T_q)] \geq b_{K_L} \cdot b_{K'_{-i}-K_L} > 0$, and so let $b^i_{K'} := b_{K_L} \cdot b_{K'_{-i}-K_L}$.

Now consider the inductive step of the nested induction, i.e., non-terminals T_i added to set $\bar{Z}_{K'}$ at step II.2. If T_i is of L-form, then for a non-terminal T_i there is a positive probability of generating a child of a non-terminal $T_j \in \bar{Z}_{K'}$, for which we already know that $\exists \sigma'_{K'} \in \Psi :\ Pr^{\sigma'_{K'}}_{T_j}[\bigcap_{q\in K'} Reach(T_q)] \geq b^j_{K'} > 0$, for some value $b^j_{K'} > 0$. Let $h := T_i(u, T_j)$. Using the strategy $\sigma'_{K'}$ in the next generation as if the play starts in it, we get an augmented strategy $\sigma'_{K'}$, such that $Pr^{\sigma'_{K'}}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] \geq p_{i,j} \cdot Pr^{\sigma'_{K'}}_{h}[\bigcap_{q\in K'} Reach(T_q)] = p_{i,j} \cdot Pr^{\sigma'_{K'}}_{T_j}[\bigcap_{q\in K'} Reach(T_q)] \geq p_{i,j} \cdot b^j_{K'} > 0$, where $p_{i,j} > 0$ is the probability of the rule $T_i \xrightarrow{p_{i,j}} T_j$. Let $b^i_{K'} := p_{i,j} \cdot b^j_{K'}$.

If T_i is of M-form, then $\exists a' \in \Gamma^i :\ T_i \stackrel{a'}{\to} T_j$, $T_j \in \bar{Z}_{K'}$, where $\exists \sigma'_{K'} \in \Psi$ such that $Pr^{\sigma'_{K'}}_{T_j}[\bigcap_{q\in K'} Reach(T_q)] \geq b^j_{K'} > 0$, for some value $b^j_{K'} > 0$. Again let $h := T_i(u, T_j)$. Hence, by combining the witness strategy $\sigma'_{K'}$ for a starting non-terminal T_j (from the nested induction assumption) with the initial local choice of choosing deterministically action a' starting at a non-terminal T_i, we obtain an augmented strategy $\sigma'_{K'}$ for a starting non-terminal T_i, such that $Pr^{\sigma'_{K'}}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = Pr^{\sigma'_{K'}}_{h}[\bigcap_{q\in K'} Reach(T_q)] = Pr^{\sigma'_{K'}}_{T_j}[\bigcap_{q\in K'} Reach(T_q)] \geq b^i_{K'} > 0$, where let $b^i_{K'} := b^j_{K'}$.

If T_i is of Q-form (i.e., $T_i \xrightarrow{1} T_j\ T_r$), then $T_j \in \bar{Z}_{K'} \vee T_r \in \bar{Z}_{K'}$, and so $\exists \sigma'_{K'} \in \Psi$: $Pr^{\sigma'_{K'}}_{T_y}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \leq 1 - b^y_{K'} < 1$, for some value $b^y_{K'} > 0$, where $y \in \{j,r\}$. Let $h_y := T_i(x, T_y)$ and $h_{\bar{y}} := T_i(\bar{x}, T_{\bar{y}})$, where $\bar{y} \in \{j,r\} - \{y\}$, $x \in \{l,r\}$ and $\bar{x} \in \{l,r\} - \{x\}$. By augmenting this $\sigma'_{K'}$ to be used from the next generation from the child of non-terminal T_y as if the play starts in it and using an arbitrary strategy from the child of non-terminal $T_{\bar{y}}$, it holds that $Pr^{\sigma'_{K'}}_{T_i}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \leq Pr^{\sigma'_{K'}}_{h_y}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \cdot Pr^{\sigma'_{K'}}_{h_{\bar{y}}}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \leq Pr^{\sigma'_{K'}}_{T_y}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \leq 1 - b^i_{K'} < 1$, where let $b^i_{K'} := b^y_{K'}$.

Finally, let $b_{K'} := \min_{T_i \in \bar{Z}_{K'}}\{b^i_{K'}\}$.

Clearly, the constructed non-static strategy $\sigma'_{K'}$ can be described in time $4^k \cdot |\mathcal{A}|^{O(1)}$.

Secondly, let us show the opposite direction, i.e., where if non-terminal $T_i \in Z_{K'}$, then $\forall \sigma \in \Psi :\ Pr^{\sigma}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = 0$. For all non-terminals $T_i \in Z_{K'}$, for a copy of non-terminal T_i in the play, it holds that: if T_i is of L-form, only a child of a non-terminal in set $Z_{K'}$ can be generated; if T_i is of M-form, regardless of player's choice on actions Γ^i, similarly only a child of a non-terminal in set $Z_{K'}$ is generated as an offspring; if T_i is of Q-form, both children have non-terminals belonging to set $Z_{K'}$. This is due to non-terminals $T_i \in Z_{K'}$ not being added to set $\bar{Z}_{K'}$ at step II.2.

Fix an arbitrary strategy σ for the player. Then starting at a non-terminal $T_i \in Z_{K'}$ and under σ, the generated play can contain only copies of non-terminals in set $Z_{K'}$, i.e., the play stays confined to non-terminals from set $Z_{K'}$ (note that the play may terminate). What is more, there is *no* Q-form non-terminal T_i in $Z_{K'}$ (whether T_i is a target from set K' or not) such that non-terminal T_i splits the job, of reaching the target non-terminals from set K', amongst its two children. In other words, for each Q-form non-terminal $T_i \in Z_{K'}$ (i.e., $T_i \xrightarrow{1} T_j\ T_r$), $\forall K_L \subset K'$ (where $K_L \neq \emptyset$): $T_j \in Z_{K_L} \vee T_r \in Z_{K'-K_L}$; and if T_i happens to be a target non-terminal itself from set K' (i.e., $i \in K'$), then $\forall K_L \subseteq K'_{-i}$: $T_j \in Z_{K_L} \vee T_r \in Z_{K'_{-i}-K_L}$ (this is due to non-terminal T_i not added to set $\bar{Z}_{K'}$ at step II.1.). So the only possibility, under σ and starting at some non-terminal $T_i \in Z_{K'}$, to generate with a positive probability a tree (play) that contains copies of all targets from set K', is (1) if all target non-terminals from set K' were never added to set $\bar{Z}_{K'}$ and, thus, belong to set $Z_{K'}$, and (2) if it is, in fact, some path w (starting at the root) in the generated tree that contains copies of all the target non-terminals from set K'. Consider such a path w and the very first copy o of any of the target non-terminals T_q ($q \in K'$) along path w. Let o be of a L-form target non-terminal T_v, let o' be the successor child of o along the path w (say of some non-terminal T_j), and let h be the ancestor history that follows along path w up until (and including) o' and ends in o' (i.e., $\texttt{current}(h) = T_j$). Then it follows that $Pr^{\sigma}_{h}[\bigcap_{q\in K'_{-v}} Reach(T_q)] > 0$. But it is easy to see that from σ one can easily construct a strategy $\sigma'_{K'_{-v}}$ such that $Pr^{\sigma'_{K'_{-v}}}_{T_j}[\bigcap_{q\in K'_{-v}} Reach(T_q)] > 0$, i.e., $T_j \in \bar{Z}_{K'_{-v}}$. But this contradicts the fact that the L-form non-terminal T_v hasn't been added to set $\bar{Z}_{K'}$ at step II.1. Similarly follows the argument for if T_v is of M-form or Q-form.

So for all non-terminals $T_i \in Z_{K'}$, regardless of strategy σ for the player, there is a zero probability of generating a play that contains all target non-terminals from set K' (i.e., $\forall \sigma \in \Psi: \ Pr^{\sigma}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = 0$). That concludes the proof. □

4.3 Algorithm for deciding $Pr^*_{T_i}[\bigcap_{q\in K} Reach(T_q)] \stackrel{?}{=} 1$

In this section we present an algorithm for deciding, given an OBMDP, $\mathcal{A}$, given a set $K \subseteq [n]$ of $k = |K|$ target non-terminals and given a starting non-terminal T_i, whether $Pr^*_{T_i}[\bigcap_{q\in K} Reach(T_q)] = \sup_{\sigma\in\Psi} Pr^{\sigma}_{T_i}[\bigcap_{q\in K} Reach(T_q)] = 1$, i.e., the optimal probability of generating a play (tree) that contains all target non-terminals from set K is $= 1$. Recall, from Example 2.1, that there need not be a strategy for the player that achieves probability exactly 1, which is the question in the next section (almost-sure multi-target reachability). However, there may nevertheless be a sequence of strategies that achieve probabilities arbitrarily close to 1 (limit-sure multi-target reachability), and the question of the existence of such a sequence is what we address in this section. In other words, we are asking whether there exists a sequence of strategies $\langle \sigma^*_{\varepsilon_j} \mid j \in \mathbb{N} \rangle$ such that $\forall j \in \mathbb{N}$, $\varepsilon_j > \varepsilon_{j+1} > 0$ (i.e., $\lim_{j\to\infty} \varepsilon_j = 0$), and $Pr^{\sigma^*_{\varepsilon_j}}_{T_i}[\bigcap_{q\in K} Reach(T_q)] \geq 1 - \varepsilon_j$. The algorithm runs in time $4^k \cdot |\mathcal{A}|^{O(1)}$, and hence is fixed-parameter tractable with respect to k.

First, as a preprocessing step, for each subset of target non-terminals $K' \subseteq K$, we compute the set $Z_{K'} := \{T_i \in V \mid \forall \sigma \in \Psi : \ Pr^{\sigma}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = 0\}$, using Algorithm 4.1. Let us also denote by AS_q, for every $q \in K$, the set of non-terminals T_j (including the target non-terminal T_q itself) for which $Pr^*_{T_j}[Reach(T_q)] = 1$. Due to the equivalence between OBMDPs and BMDPs with respect to single-target reachability (see subsection 2.4.1), these sets can be computed in P-time by applying the algorithm

Algorithm 4.2 Algorithm for limit-sure multi-target reachability in a given OBMDP. The output is the set $F_K = \{T_i \in V \mid Pr^*_{T_i}[\bigcap_{q\in K} Reach(T_q)] = 1\}$.

I. Let $F_{\{q\}} := AS_q$ and $S_{\{q\}} := V - F_{\{q\}} - Z_{\{q\}}$, for each $q \in K$. Let $F_\emptyset := V$ and $S_\emptyset := \emptyset$.

II. For $l = 2 \dots k$:

For every subset of target non-terminals $K' \subseteq K$ of size $|K'| = l$:

1. $D_{K'} := \{T_i \in V - Z_{K'} \mid$ one of the following holds:
   - T_i is of L-form where $i \in K'$, $T_i \not\rightarrow \varnothing$ and $\forall T_j \in V$: if $T_i \rightarrow T_j$, then $T_j \in F_{K'_{-i}}$.
   - T_i is of M-form where $i \in K'$ and $\exists a^* \in \Gamma^i : T_i \xrightarrow{a^*} T_j,\ T_j \in F_{K'_{-i}}$.
   - T_i is of Q-form $(T_i \xrightarrow{1} T_j\ T_r)$ where $i \in K'$ and $\exists K_L \subseteq K'_{-i} : T_j \in F_{K_L} \wedge T_r \in F_{K'_{-i}-K_L}$.
   - T_i is of Q-form $(T_i \xrightarrow{1} T_j\ T_r)$ where $\exists K_L \subset K'$ $(K_L \neq \emptyset) : T_j \in F_{K_L} \wedge T_r \in F_{K'-K_L}$.$\}$
2. Repeat until no change has occurred to $D_{K'}$:
   (a) add $T_i \notin D_{K'}$ to $D_{K'}$, if of L-form, $T_i \not\rightarrow \varnothing$ and $\forall T_j \in V$: if $T_i \rightarrow T_j$, then $T_j \in D_{K'}$.
   (b) add $T_i \notin D_{K'}$ to $D_{K'}$, if of M-form and $\exists a^* \in \Gamma^i : T_i \xrightarrow{a^*} T_j,\ T_j \in D_{K'}$.
   (c) add $T_i \notin D_{K'}$ to $D_{K'}$, if of Q-form $(T_i \xrightarrow{1} T_j\ T_r)$ and $T_j \in D_{K'} \vee T_r \in D_{K'}$.
3. Let $X := V - (D_{K'} \cup Z_{K'})$.
4. Initialize $S_{K'} := \{T_i \in X \mid$ either $i \in K'$, or T_i is of L-form and $T_i \rightarrow \varnothing \vee T_i \rightarrow T_j,\ T_j \in Z_{K'}\} \cup \bigcup_{\emptyset \subset K'' \subset K'} (X \cap S_{K''})$.
5. Repeat until no change has occurred to $S_{K'}$:
   (a) add $T_i \in X - S_{K'}$ to $S_{K'}$, if of L-form and $T_i \rightarrow T_j,\ T_j \in S_{K'} \cup Z_{K'}$.
   (b) add $T_i \in X - S_{K'}$ to $S_{K'}$, if of M-form and $\forall a \in \Gamma^i :\ T_i \xrightarrow{a} T_j,\ T_j \in S_{K'} \cup Z_{K'}$.
   (c) add $T_i \in X - S_{K'}$ to $S_{K'}$, if of Q-form $(T_i \xrightarrow{1} T_j\ T_r)$ and $T_j \in S_{K'} \cup Z_{K'}\ \wedge\ T_r \in S_{K'} \cup Z_{K'}$.
6. $\mathcal{C} \leftarrow$ MEC decomposition of $G[X - S_{K'}]$.
7. For every $q \in K'$, let $H_q := \{T_i \in X - S_{K'} \mid T_i$ is of Q-form $(T_i \xrightarrow{1} T_j\ T_r)$ and $((T_j \in X - S_{K'} \wedge T_r \in \bar{Z}_{\{q\}}) \vee (T_j \in \bar{Z}_{\{q\}} \wedge T_r \in X - S_{K'}))\}$.
8. Let $F_{K'} := \bigcup\{C \in \mathcal{C} \mid P_C = K' \vee (P_C \neq \emptyset \wedge P_C \neq K' \wedge \exists T_i \in C, \exists a \in \Gamma^i :\ T_i \xrightarrow{a} T_j,\ T_j \in F_{K'-P_C})\}$, where $P_C = \{q \in K' \mid C \cap H_q \neq \emptyset\}$.
9. Repeat until no change has occurred to $F_{K'}$:
   (a) add $T_i \in X - (S_{K'} \cup F_{K'})$ to $F_{K'}$, if of L-form and $T_i \rightarrow T_j,\ T_j \in F_{K'} \cup D_{K'}$.
   (b) add $T_i \in X - (S_{K'} \cup F_{K'})$ to $F_{K'}$, if of M-form and $\exists a^* \in \Gamma^i : T_i \xrightarrow{a^*} T_j,\ T_j \in F_{K'}$.
   (c) add $T_i \in X - (S_{K'} \cup F_{K'})$ to $F_{K'}$, if of Q-form $(T_i \xrightarrow{1} T_j\ T_r)$ and $T_j \in F_{K'} \vee T_r \in F_{K'}$.
10. If $X \neq S_{K'} \cup F_{K'}$, let $S_{K'} := X - F_{K'}$ and go to step 5.
11. Else, i.e., if $X = S_{K'} \cup F_{K'}$, let $F_{K'} := F_{K'} \cup D_{K'}$.

III. **Output** F_K.

from [ESY18, Theorem 9.3] to each target non-terminal T_q, $q \in K$. Recall that it was shown in [ESY18, Theorem 9.4] that for (O)BMDPs with a single target the almost-sure and limit-sure reachability problems coincide. So in fact, for every $q \in K$, there exists a strategy τ_q such that for every $T_j \in AS_q$: $Pr^{\tau_q}_{T_j}[Reach(T_q)] = 1$.

After this preprocessing step, we apply Algorithm 4.2 to identify the non-terminals T_i for which $Pr^*_{T_i}[\bigcap_{q\in K} Reach(T_q)] = 1$. Again let K'_{-i} denote the set $K' - \{i\}$. Also, to recall what notation $T_i \to T_j$ or $T_i \to \varnothing$ means, refer to the paragraph before Definition 12. And for the definition of MEC and MEC-decomposition, refer to Definition 13, where recall that in our setting the partition of the dependency graph nodes, $U = V$, that we use is $U_P := \{T_i \in U \mid T_i \text{ is of L-form}\}$ and $U_1 := \{T_i \in U \mid T_i$ is of M-form or Q-form$\}$.

Theorem 4.3. *Algorithm 4.2 computes, given an OBMDP, $\mathcal{A}$, and a set $K \subseteq [n]$ of $k = |K|$ target non-terminals, for each subset $K' \subseteq K$, the set of non-terminals $F_{K'} := \{T_i \in V \mid Pr^*_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = 1\}$. The algorithm runs in time $4^k \cdot |\mathcal{A}|^{O(1)}$. Moreover, for each $K' \subseteq K$, given $\varepsilon > 0$, the algorithm can also be augmented to compute a randomized non-static strategy $\sigma^{\varepsilon}_{K'}$ such that $Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] \geq 1 - \varepsilon$ for all non-terminals $T_i \in F_{K'}$.*

Proof. We will refer to the loop executing steps II.5. through II.10. for a specific subset $K' \subseteq K$ as the "inner" loop and the iteration through all subsets of K as the "outer" loop. Clearly the inner loop terminates, due to step II.10. always adding at least one non-terminal to set $S_{K'}$ and step II.11. eventually executing. The running time of the algorithm follows from the facts that the outer loop executes for 2^k iterations and inside each iteration of the outer loop, steps II.1. and II.4. require time at most $2^k \cdot |\mathcal{A}|^{O(1)}$ and the inner loop executes for at most $|V|$ iterations, where during each inner loop iteration the steps in it execute in time at most $|\mathcal{A}|^{O(1)}$.

For the proof of correctness, we show that for every subset of target non-terminals $K' \subseteq K$, $F_{K'}$ (from the decomposition $V = F_{K'} \cup S_{K'} \cup Z_{K'}$) is the set of non-terminals T_i for which the following property holds:

$$(A)^i_{K'}\text{: } \sup_{\sigma\in\Psi} Pr^{\sigma}_{T_i}[\textstyle\bigcap_{q\in K'} Reach(T_q)] = Pr^*_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = 1, \text{ i.e.,}$$
$$\forall \varepsilon > 0,\ \exists \sigma^{\varepsilon}_{K'} \in \Psi \text{ such that } Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}[\textstyle\bigcap_{q\in K'} Reach(T_q)] \geq 1 - \varepsilon.$$

Otherwise, if $T_i \in S_{K'}$, then we show that the following property holds:

$$(B)^i_{K'}\text{: } \sup_{\sigma\in\Psi} Pr^{\sigma}_{T_i}[\textstyle\bigcap_{q\in K'} Reach(T_q)] < 1, \text{ i.e., there exists a value } g > 0 \text{ such that}$$
$$\forall \sigma \in \Psi :\ Pr^{\sigma}_{T_i}[\textstyle\bigcap_{q\in K'} Reach(T_q)] \leq 1 - g.$$

Clearly, for non-terminals $T_i \in Z_{K'}$, property $(B)^i_{K'}$ is satisfied, since it holds that $\sup_{\sigma\in\Psi} Pr^{\sigma}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = 0 < 1$ (by Proposition 4.2). Finally, the answer for the full set of targets is $F := F_K$.

We base this proof on an induction on the size of subset K', i.e., on the time of computing sets $S_{K'}$ and $F_{K'}$ for $K' \subseteq K$. And in the process, for each subset $K' \subseteq K$ of target non-terminals, we show how to construct a randomized *non-static* strategy $\sigma^{\varepsilon}_{K'}$ (for any given $\varepsilon > 0$) that ensures $Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] \geq 1-\varepsilon$ for each non-terminal $T_i \in F_{K'}$.

Clearly for any subset of target non-terminals, $K' := \{q\} \subseteq K$, of size $l = 1$, each non-terminal $T_i \in F_{\{q\}}$ (respectively, $T_i \in V - F_{\{q\}}$) satisfies property $(A)^i_{\{q\}}$ (respectively, $(B)^i_{\{q\}}$), due to step I. and the definition of the $AS_q, q \in K$ sets. Furthermore, for each such subset $\{q\} \subseteq K$, there is in fact a strategy $\sigma_{\{q\}}$ such that $\forall T_i \in F_{\{q\}}$: $Pr^{\sigma_{\{q\}}}_{T_i}[Reach(T_q)] = 1$. Moreover, by [ESY18, Theorem 9.4], this strategy $\sigma_{\{q\}}$ is non-static and deterministic. Analysing subset K' of target non-terminals of size l as part of step II., assume that, for every $K'' \subset K'$ of size $\leq l-1$, sets $S_{K''}$ and $F_{K''}$ have already been computed, and for each non-terminal T_j belonging to set $F_{K''}$ (respectively, set $S_{K''}$) property $(A)^j_{K''}$ (respectively, $(B)^j_{K''}$) holds. That is, by induction assumption, for each $K'' \subset K'$, for every $\varepsilon > 0$ there is a randomized non-static strategy $\sigma^{\varepsilon}_{K''}$ such that for any $T_j \in F_{K''}$: $Pr^{\sigma^{\varepsilon}_{K''}}_{T_j}[\bigcap_{q\in K''} Reach(T_q)] \geq 1-\varepsilon$, and also for any $T_j \in S_{K''}$: $\sup_{\sigma\in\Psi} Pr^{\sigma}_{T_j}[\bigcap_{q\in K''} Reach(T_q)] < 1$. We now need to show that at end of the inner loop analysis of subset K', property $(A)^i_{K'}$ (respectively, $(B)^i_{K'}$) holds for every non-terminal $T_i \in F_{K'}$ (respectively, $T_i \in S_{K'}$).

First we show that property $(A)^i_{K'}$ holds for each non-terminal T_i belonging to set $D_{K'}$ ($\subseteq F_{K'}$), precomputed prior to the execution of the inner loop for K'.

Lemma 4.4. *Every non-terminal $T_i \in D_{K'}$ satisfies property $(A)^i_{K'}$.*

Proof. The lemma is proved via a nested induction based on the time when a non-terminal is added to set $D_{K'}$. Consider the base case where $T_i \in D_{K'}$ is a non-terminal, added at the initialization step II.1.

(i) Suppose T_i is of L-form where $i \in K'$ and for all associated rules a child is generated that is of a non-terminal $T_j \in F_{K'_{-i}}$, where property $(A)^j_{K'_{-i}}$ holds. Then, for every $\varepsilon > 0$, using the witness strategy $\sigma^{\varepsilon}_{K'_{-i}}$ from the induction assumption for all such non-terminals T_j in the next generation, as if the play starts in it, we obtain a strategy $\sigma^{\varepsilon}_{K'}$ for a starting (target) non-terminal T_i such that:

$$Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}\Big[\bigcap_{q\in K'} Reach(T_q)\Big] = Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}\Big[\bigcap_{q\in K'_{-i}} Reach(T_q)\ \Big|\ Reach(T_i)\Big]\cdot Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}\Big[Reach(T_i)\Big]$$

$$= Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}\Big[\bigcap_{q\in K'_{-i}} Reach(T_q)\Big] = \sum_j p_{i,j}\cdot Pr^{\sigma^{\varepsilon}_{K'}}_{T_i(u,T_j)}\Big[\bigcap_{q\in K'_{-i}} Reach(T_q)\Big]$$

$$= \sum_j p_{i,j}\cdot Pr^{\sigma^{\varepsilon}_{K'_{-i}}}_{T_j}\Big[\bigcap_{q\in K'_{-i}} Reach(T_q)\Big] \geq \sum_j p_{i,j}\cdot(1-\varepsilon) = 1-\varepsilon$$

where $p_{i,j} > 0$ is the probability of rule $T_i \xrightarrow{p_{i,j}} T_j$.

(ii) Suppose T_i is of M-form where $i\in K'$ and $\exists a^*\in\Gamma^i$ such that $T_i \xrightarrow{a^*} T_j$, $T_j \in F_{K'_{-i}}$, where property $(A)^j_{K'_{-i}}$ holds. Let $h := T_i(u,T_j)$. By combining the witness strategies $\sigma^{\varepsilon}_{K'_{-i}}$, for every $\varepsilon > 0$, from property $(A)^j_{K'_{-i}}$ from the induction assumption for non-terminal T_j, as if the play starts in it, with the initial local choice of choosing action a^* deterministically starting at a non-terminal T_i, we obtain for every $\varepsilon > 0$ a combined strategy $\sigma^{\varepsilon}_{K'}$ such that starting at a (target) non-terminal T_i, it follows that $Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = Pr^{\sigma^{\varepsilon}_{K'}}_{h}[\bigcap_{q\in K'_{-i}} Reach(T_q)] = Pr^{\sigma^{\varepsilon}_{K'_{-i}}}_{T_j}[\bigcap_{q\in K'_{-i}} Reach(T_q)] \geq 1-\varepsilon$.

(iii) Suppose T_i is of Q-form ($T_i \xrightarrow{1} T_j\ T_r$) where $i\in K'$ and there exists a split of the rest of the target non-terminals, implied by $K_L \subseteq K'_{-i}$ and $K'_{-i} - K_L$, such that $T_j \in F_{K_L} \wedge T_r \in F_{K'_{-i}-K_L}$. Let $h_l := T_i(l,T_j)$ and $h_r := T_i(r,T_r)$. For every $\varepsilon > 0$, if we let $\varepsilon' := 1-\sqrt{1-\varepsilon}$, then by combining the two witness strategies $\sigma^{\varepsilon'}_{K_L}$ and $\sigma^{\varepsilon'}_{K'_{-i}-K_L}$ from the induction assumption for non-terminals T_j and T_r, respectively, to be used in the next generation as if the play starts in it, we obtain a strategy $\sigma^{\varepsilon}_{K'}$ for a starting (target) non-terminal T_i such that $Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}[\bigcap_{q\in K'_{-i}} Reach(T_q)] \geq Pr^{\sigma^{\varepsilon}_{K'}}_{h_l}[\bigcap_{q\in K_L} Reach(T_q)]\cdot Pr^{\sigma^{\varepsilon}_{K'}}_{h_r}[\bigcap_{q\in K'_{-i}-K_L} Reach(T_q)] = Pr^{\sigma^{\varepsilon'}_{K_L}}_{T_j}[\bigcap_{q\in K_L} Reach(T_q)]\cdot Pr^{\sigma^{\varepsilon'}_{K'_{-i}-K_L}}_{T_r}[\bigcap_{q\in K'_{-i}-K_L} Reach(T_q)] \geq (1-\varepsilon')^2 = (\sqrt{1-\varepsilon})^2 = 1-\varepsilon$.

(iv) Suppose T_i is of Q-form ($T_i \xrightarrow{1} T_j\ T_r$) where there exists a proper split of the target non-terminals from set K', implied by $K_L \subset K'$ (where $K_L \neq \emptyset$) and $K' - K_L$, such that $T_j \in F_{K_L} \wedge T_r \in F_{K'-K_L}$. Similarly, for every $\varepsilon > 0$, let $\varepsilon' := 1-\sqrt{1-\varepsilon}$ and combine the two witness strategies $\sigma^{\varepsilon'}_{K_L}$ and $\sigma^{\varepsilon'}_{K'-K_L}$ from the induction assumption for non-terminals T_j and T_r in the same way as in (iii). It follows that property $(A)^i_{K'}$ is satisfied.

Now consider non-terminals T_i added to set $D_{K'}$ at step II.2. If T_i is of L-form, then all associated rules generate children of non-terminals T_j already in set $D_{K'}$, where $(A)^j_{K'}$ holds by the (nested) induction. So using, for every $\varepsilon > 0$, the strategy $\sigma^\varepsilon_{K'}$ from the nested induction assumption for all such non-terminal T_j in the next generation, as if the play starts in it, and applying the same argument as in (i), then property $(A)^i_{K'}$ is also satisfied.

If T_i is of M-form, then $\exists a^* \in \Gamma^i : T_i \xrightarrow{a^*} T_j,\ T_j \in D_{K'}$. Again let $h := T_i(u, T_j)$. By combining, for every $\varepsilon > 0$, the witness strategy $\sigma^\varepsilon_{K'}$ for non-terminal T_j (from the nested induction assumption), as if the play starts in it, with the initial local choice of choosing action a^* deterministically starting at a non-terminal T_i, we obtain an augmented strategy $\sigma^\varepsilon_{K'}$ for a starting non-terminal T_i such that $Pr^{\sigma^\varepsilon_{K'}}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = Pr^{\sigma^\varepsilon_{K'}}_{h}[\bigcap_{q\in K'} Reach(T_q)] = Pr^{\sigma^\varepsilon_{K'}}_{T_j}[\bigcap_{q\in K'} Reach(T_q)] \geq 1-\varepsilon$.

If T_i is of Q-form ($T_i \xrightarrow{1} T_j\ T_r$), then $T_j \in D_{K'} \vee T_r \in D_{K'}$, i.e., for every $\varepsilon > 0$, $\exists \sigma^\varepsilon_{K'} \in \Psi$ such that $Pr^{\sigma^\varepsilon_{K'}}_{T_y}[\bigcap_{q\in K'} Reach(T_q)] \geq 1-\varepsilon \Leftrightarrow Pr^{\sigma^\varepsilon_{K'}}_{T_y}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \leq \varepsilon$, where $y \in \{j, r\}$. Let $h_y := T_i(x, T_y)$ and $h_{\bar{y}} := T_i(\bar{x}, T_{\bar{y}})$, where $\bar{y} \in \{j, r\} - \{y\}$, $x \in \{l, r\}$ and $\bar{x} \in \{l, r\} - \{x\}$. By augmenting strategy $\sigma^\varepsilon_{K'}$ to be used from the next generation from the child of non-terminal T_y, as if the play starts in it, and using an arbitrary strategy from the child of non-terminal $T_{\bar{y}}$, it follows that $Pr^{\sigma^\varepsilon_{K'}}_{T_i}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \leq Pr^{\sigma^\varepsilon_{K'}}_{h_y}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \cdot Pr^{\sigma^\varepsilon_{K'}}_{h_{\bar{y}}}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \leq Pr^{\sigma^\varepsilon_{K'}}_{T_y}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \leq \varepsilon \Leftrightarrow Pr^{\sigma^\varepsilon_{K'}}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] \geq 1-\varepsilon$, i.e., property $(A)^i_{K'}$ holds. □

Next, we show that if $T_i \in S_{K'}$, then property $(B)^i_{K'}$ is satisfied.

Lemma 4.5. *Every non-terminal $T_i \in S_{K'}$ satisfies property $(B)^i_{K'}$.*

Proof. Again this is proved via a nested induction based on the time a non-terminal is added to set $S_{K'}$. Assuming that all non-terminals T_j, added already to set $S_{K'}$ in previous steps and iterations of the inner loop, satisfy $(B)^j_{K'}$, then we need to show that for a new addition T_i to set $S_{K'}$, property $(B)^i_{K'}$ also holds.

Consider the non-terminals T_i added to set $S_{K'}$ at the initialization step II.4.

If T_i is of L-form where $T_i \to \varnothing \vee T_i \to T_j,\ T_j \in Z_{K'}$, then with a constant positive probability non-terminal T_i immediately either does not generate any offspring at all or generates a child of non-terminal $T_j \in Z_{K'}$, for which we already know that $(B)^j_{K'}$ holds. It is clear that property $(B)^i_{K'}$ is also satisfied.

If, for some subset $K'' \subset K'$, $T_i \in S_{K''}$, i.e., property $(B)^i_{K''}$ holds, then there is a value $g > 0$ such that $\forall \sigma \in \Psi :\ Pr^\sigma_{T_i}[\bigcap_{q\in K'} Reach(T_q)] \leq Pr^\sigma_{T_i}[\bigcap_{q\in K''} Reach(T_q)] \leq 1-g$

and so property $(B)^i_{K'}$ is also satisfied. Note that if, for some subset $K'' \subset K'$, $T_i \in Z_{K''}$, then similarly $T_i \in Z_{K'}$ and so already $T_i \notin X$.

If T_i is a target non-terminal in set K' (i.e., $i \in K'$), then since it has not been added to set $D_{K'}$ in step II.1: (1) if of L-form, it generates with a constant positive probability a child of non-terminal $T_j \in S_{K'_{-i}} \cup Z_{K'_{-i}}$, where $(B)^j_{K'_{-i}}$ holds; (2) if of M-form, irrespective of the strategy it generates a child of non-terminal $T_j \in S_{K'_{-i}} \cup Z_{K'_{-i}}$, where again $(B)^j_{K'_{-i}}$ holds; (3) and if of Q-form, it generates two children of non-terminals T_j, T_r, for which no matter how we split the rest of the target non-terminals from set K'_{-i} (into subsets $K_L \subseteq K'_{-i}$ and $K'_{-i} - K_L$), either $(B)^j_{K_L}$ holds or $(B)^r_{K'_{-i}-K_L}$ holds. In other words, for a target non-terminal T_i in the initial set $S_{K'}$ there is no sequence of strategies to ensure that the rest of the target non-terminals are reached with probability arbitrarily close to 1 (the reasoning behind this last statement is the same as the arguments in (i) - (iii) below, since for a starting (target) non-terminal T_i: $\forall \sigma \in \Psi: \ Pr^\sigma_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = Pr^\sigma_{T_i}[\bigcap_{q\in K'_{-i}} Reach(T_q)]$).

Observe that by the end of step II.4. all target non-terminals T_q ($q \in K'$) belong either to set $D_{K'}$ or set $S_{K'}$. Now consider a non-terminal T_i added to set $S_{K'}$ in step II.5. during some iteration of the inner loop.

(i) Suppose T_i is of L-form. Then $T_i \to T_j$, $T_j \in S_{K'} \cup Z_{K'}$, where $(B)^j_{K'}$ holds. So irrespective of the strategy there is a constant positive probability to generate a child of the above non-terminal T_j such that $Pr^*_{T_j}[\bigcap_{q\in K'} Reach(T_q)] < 1$, or in other words, $\exists g > 0$ such that $\forall \sigma \in \Psi: \ Pr^\sigma_{T_j}[\bigcap_{q\in K'} Reach(T_q)] \leq 1 - g$. Let $h := T_i(u, T_j)$. But there is a value $g > 0$ such that $\forall \sigma \in \Psi: \ Pr^\sigma_{T_i}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \geq p_{i,j} \cdot Pr^\sigma_h[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \geq p_{i,j} \cdot g$ if and only if $\forall \sigma \in \Psi: \ Pr^\sigma_{T_j}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \geq g$, where $p_{i,j} > 0$ is the probability of the rule $T_i \xrightarrow{p_{i,j}} T_j$. And since the latter part of the statement holds, then the former, showing property $(B)^i_{K'}$, also holds.

(ii) Suppose T_i is of M-form. Then $\forall a \in \Gamma^i: \ T_i \xrightarrow{a} T_j$, $T_j \in S_{K'} \cup Z_{K'}$. So irrelevant of strategy σ, starting in a non-terminal T_i the next generation surely consists of some non-terminal T_j satisfying property $\sup_{\sigma\in\Psi} Pr^\sigma_{T_j}[\bigcap_{q\in K'} Reach(T_q)] < 1$, i.e., $\forall \sigma \in \Psi: \ Pr^\sigma_{T_j}[\bigcap_{q\in K'} Reach(T_q)] \leq 1 - g$, for some value $g > 0$. Clearly, for some value $g > 0$, $\forall \sigma \in \Psi: \ Pr^\sigma_{T_i}[\bigcap_{q\in K'} Reach(T_q)] \leq \max_{\{T_j \in S_{K'} \cup Z_{K'} | T_i \to T_j\}} Pr^\sigma_{T_i(u,T_j)}[\bigcap_{q\in K'} Reach(T_q)] \leq 1 - g$ (i.e., property $(B)^i_{K'}$) if and only if $\forall \sigma \in \Psi$: $\max_{\{T_j \in S_{K'} \cup Z_{K'} | T_i \to T_j\}} Pr^\sigma_{T_j}[\bigcap_{q\in K'} Reach(T_q)] \leq 1 - g$, where the latter is satisfied.

(iii) Suppose T_i is of Q-form (i.e., $T_i \xrightarrow{1} T_j\ T_r$), then $T_j, T_r \in S_{K'} \cup Z_{K'}$, i.e., both $(B)^j_{K'}$ and $(B)^r_{K'}$ are satisfied. We know that:

1) Neither of the two children can single-handedly reach all target non-terminals from set K' with probability arbitrarily close to 1. That is, for some value $g > 0, \forall \sigma \in \Psi$: $Pr^{\sigma}_{T_j}[\bigcap_{q \in K'} Reach(T_q)] \leq 1 - g$ and $Pr^{\sigma}_{T_r}[\bigcap_{q \in K'} Reach(T_q)] \leq 1 - g$.

2) Moreover, since T_i was not added to set $D_{K'}$ in step II.1., then $\forall K_L \subset K'$ (where $K_L \neq \emptyset$) either $(B)^j_{K_L}$ holds (i.e., $T_j \notin F_{K_L}$) or $(B)^r_{K'-K_L}$ holds (i.e., $T_r \notin F_{K'-K_L}$), i.e., there is some value $g > 0$ such that either $\forall \sigma \in \Psi$: $Pr^{\sigma}_{T_j}[\bigcap_{q \in K_L} Reach(T_q)] \leq 1 - g$ or $\forall \sigma \in \Psi$: $Pr^{\sigma}_{T_r}[\bigcap_{q \in K'-K_L} Reach(T_q)] \leq 1 - g$.

(Statements 1) and 2) hold for the same value $g > 0$, since there are only finitely many subsets of K', so we can take g to be the minimum of all such values from all the properties $(B)^{j/r}_{K''}$ $(K'' \subseteq K')$.)

Let $h_l := T_i(l, T_j)$ and $h_r := T_i(r, T_r)$. Notice that for any strategy $\sigma \in \Psi$ and for any $q' \in K'$, $Pr^{\sigma}_{T_i}[\bigcup_{q \in K'} Reach^{\complement}(T_q)] \geq Pr^{\sigma}_{T_i}[Reach^{\complement}(T_{q'})] = Pr^{\sigma}_{h_l}[Reach^{\complement}(T_{q'})] \cdot Pr^{\sigma}_{h_r}[Reach^{\complement}(T_{q'})]$.

We claim that there is a value $g_i > 0$ such that $\forall \sigma \in \Psi$: $\bigvee_{q \in K'} Pr^{\sigma}_{T_j}[Reach^{\complement}(T_q)] \cdot Pr^{\sigma}_{T_r}[Reach^{\complement}(T_q)] \geq g_i$. But for any $q \in K'$ and for any $\sigma \in \Psi$ one can obviously construct $\sigma' \in \Psi$ such that $Pr^{\sigma}_{T_j}[Reach^{\complement}(T_q)] = Pr^{\sigma'}_{h_l}[Reach^{\complement}(T_q)]$ and similarly for non-terminal T_r. Therefore, it follows from the claim that $\forall \sigma \in \Psi$: $\bigvee_{q \in K'} Pr^{\sigma}_{h_l}[Reach^{\complement}(T_q)] \cdot Pr^{\sigma}_{h_r}[Reach^{\complement}(T_q)] \geq g_i$ and, therefore, it follows that $\forall \sigma \in \Psi$: $Pr^{\sigma}_{T_i}[\bigcup_{q \in K'} Reach^{\complement}(T_q)] \geq g_i \Leftrightarrow Pr^{\sigma}_{T_i}[\bigcap_{q \in K'} Reach(T_q)] \leq 1 - g_i$.

Suppose the opposite, i.e., assume $(\mathcal{P})$ that $\forall g' > 0$, $\exists \sigma_{g'} \in \Psi$ such that $\bigwedge_{q \in K'} Pr^{\sigma_{g'}}_{T_j}[Reach^{\complement}(T_q)] \cdot Pr^{\sigma_{g'}}_{T_r}[Reach^{\complement}(T_q)] < g'$. Now for any $q \in K'$, by statement 2) above, we know that $T_j \notin F_{\{q\}} \vee T_r \notin F_{K'_{-q}}$ and $T_j \notin F_{K'_{-q}} \vee T_r \notin F_{\{q\}}$. First, suppose that in fact for some $q' \in K'$ it is the case that $T_j \notin F_{\{q'\}} \wedge T_r \notin F_{\{q'\}}$ (i.e., $T_j \in S_{\{q'\}} \cup Z_{\{q'\}} \wedge T_r \in S_{\{q'\}} \cup Z_{\{q'\}}$). That is, for some value $g > 0$, $\forall \sigma \in \Psi$: $Pr^{\sigma}_{T_j}[Reach^{\complement}(T_{q'})] \geq g$ and $Pr^{\sigma}_{T_r}[Reach^{\complement}(T_{q'})] \geq g$, where our claim follows directly by letting $g_i := g^2$ (hence, contradiction to $(\mathcal{P})$). Second, suppose that for some $q' \in K'$ it is the case that $T_j \notin F_{K'_{-q'}} \wedge T_r \notin F_{K'_{-q'}}$ (i.e., $T_j \in S_{K'_{-q'}} \cup Z_{K'_{-q'}} \wedge T_r \in S_{K'_{-q'}} \cup Z_{K'_{-q'}}$). But then T_i would have been added to set $S_{K'_{-q'}}$ at step II.5.(c) when constructing the answer for subset of targets $K'_{-q'}$.

However, we already know that $T_i \in \bigcap_{K''\subset K'} F_{K''}$ (following from steps II.3. and II.4. that $T_i \notin \bigcup_{K''\subset K'}(S_{K''} \cup Z_{K''})$). Hence, again a contradiction.

Therefore, it follows that for every $q \in K'$, either $T_j \notin F_{\{q\}} \wedge T_j \notin F_{K'_{-q}}$ or $T_r \notin F_{\{q\}} \wedge T_r \notin F_{K'_{-q}}$. And in particular, the essential part is that $\forall q \in K'$, either $T_j \notin F_{\{q\}}$ or $T_r \notin F_{\{q\}}$. That is, for every $q \in K'$, for some value $g > 0$ either $\forall\sigma \in \Psi : Pr^{\sigma}_{T_j}[Reach^{\complement}(T_q)] \geq g$, or $\forall\sigma \in \Psi : Pr^{\sigma}_{T_r}[Reach^{\complement}(T_q)] \geq g$. But then, combined with assumption $(\mathcal{P})$, it actually follows that there exists a subset $K'' \subseteq K'$ such that $\forall\varepsilon > 0$, $\exists\sigma_\varepsilon \in \Psi$ such that $\bigwedge_{q\in K''} Pr^{\sigma_\varepsilon}_{T_r}[Reach^{\complement}(T_q)] \leq \varepsilon \wedge \bigwedge_{q\in K'-K''} Pr^{\sigma_\varepsilon}_{T_j}[Reach^{\complement}(T_q)] \leq \varepsilon$. And by Proposition 2.2(5.), it follows that $\forall\varepsilon > 0$, $\exists\sigma'_\varepsilon \in \Psi :\ Pr^{\sigma'_\varepsilon}_{T_r}[\bigcap_{q\in K''} Reach(T_q)] \geq 1-\varepsilon \ \wedge\ Pr^{\sigma'_\varepsilon}_{T_j}[\bigcap_{q\in K'-K''} Reach(T_q)] \geq 1-\varepsilon$, i.e., $T_j \in F_{K'-K''} \wedge T_r \in F_{K''}$, contradicting the known facts 1) and 2). Hence, assumption $(\mathcal{P})$ is wrong and our claim is satisfied.

Now consider non-terminals T_i added to set $S_{K'}$ in step II.10. at some iteration of the inner loop, i.e., $T_i \in Y_{K'} := X - (S_{K'} \cup F_{K'}) \subseteq \bar{Z}_{K'}$. Due to the fact that T_i has not been added previously to sets $D_{K'}$, $S_{K'}$ or $F_{K'}$, then all of the following hold:

(1.) $i \notin K'$;

(2.) if T_i is of L-form, then a non-terminal T_i generates with probability 1 a non-terminal which belongs to set $Y_{K'}$ (otherwise T_i would have been added to sets $S_{K'}$ or $F_{K'}$ in step II.4., II.5. or step II.9., respectively);

(3.) if T_i is of M-form, then $\forall a \in \Gamma^i :\ T_i \xrightarrow{a} T_d,\ T_d \notin F_{K'} \cup D_{K'}$ (otherwise T_i would have been added to sets $F_{K'}$ or $D_{K'}$ in step II.2. or step II.9., respectively), and $\exists' a \in \Gamma^i :\ T_i \xrightarrow{a'} T_j,\ T_j \notin S_{K'} \cup Z_{K'}$, i.e., $T_j \in Y_{K'}$ (otherwise T_i would have been added to set $S_{K'}$ in step II.5.); and

(4.) if T_i is of Q-form ($T_i \xrightarrow{1} T_j\ T_r$), then w.l.o.g. $T_j \in Y_{K'}$ and $T_r \in Y_{K'} \cup S_{K'} \cup Z_{K'}$ (since T_i has not been added to the other sets in steps II.2., II.5., or II.9.).

Observe that any MEC in subgraph $G[X - S_{K'}]$, that contains a node from set $Y_{K'}$, is in fact entirely contained in subgraph $G[Y_{K'}]$, and also that there is at least one MEC in $G[Y_{K'}]$. This is due to statements (2.) - (4.) and the two key facts that all nodes in $G[Y_{K'}]$ have at least one outgoing edge and there is only a finite number of nodes.

However, consider any MEC, C, in $G[Y_{K'}]$ ($Y_{K'} \subseteq X - S_{K'}$). As C has not been added to set $F_{K'}$ at step II.8., then $P_C \neq K'$ (where $P_C = \{q \in K' \mid C \cap H_q \neq \emptyset\}$) and:

- either $P_C = \emptyset$,
- or $P_C \neq \emptyset$ and for every $T_u \in C$ of M-form it holds that $\forall b \in \Gamma^u : T_u \xrightarrow{b} T_v, T_v \notin F_{K'-P_C}$.

First, let us focus on the second point. Note that for any non-terminal $T_j \in C$, clearly $T_j \in F_{P_C}$, and in fact, $\exists \sigma_{P_C} \in \Psi : Pr_{T_j}^{\sigma_{P_C}}[\bigcap_{q \in P_C} Reach(T_q)] = 1$. That is because, starting at a non-terminal $T_j \in C$, due to C being a MEC in $G[Y_{K'}]$, such a strategy σ_{P_C} can ensure that, for each $q \in P_C$, infinitely often a copy of a Q-form non-terminal in set $H_q \cap C$ is generated, which in turn spawns an independent copy of some non-terminal in set $\bar{Z}_{\{q\}}$ and thus infinitely often provides a positive probability bounded away from zero (by Proposition 4.2) to reach target non-terminal T_q.

(*) We claim that for any Q-form non-terminal $T_i \in C$ (i.e., $T_i \xrightarrow{1} T_j\, T_r$ where w.l.o.g. $T_j \in C \subseteq Y_{K'}$), it is guaranteed that $T_r \notin F_{K'-P_C}$. To see this, if it was the case that $T_r \in F_{K'-P_C}$, then, since $T_j \in F_{P_C}$, it would follow that T_i would have been added to set $D_{K'}$ in step II.1., leading to a contradiction.

(**) What is more, due to the definition of set P_C, it follows that for any Q-form non-terminal $T_i \in C$ (i.e., $T_i \xrightarrow{1} T_j\, T_r$ where w.l.o.g. $T_j \in C$), $T_r \in \bigcap_{q' \in K'-P_C} Z_{\{q'\}}$, i.e., $\sup_{\sigma \in \Psi} Pr_{T_r}^{\sigma}[Reach(T_{q'})] = 0$, for each $q' \in K' - P_C$. Note also that $T_r \notin C$, since $C \subseteq Y_{K'} \subseteq \bar{Z}_{K'} \subseteq \bar{Z}_{\{q\}}$, $\forall q \in K'$ (so if $T_r \in C$, then $P_C = K'$ and C would have been added to set $F_{K'}$ in step II.8.).

Note that property (**) implies property (*), because by the definition of the F and Z sets, if $T_r \in \bigcap_{q' \in K'-P_C} Z_{\{q'\}}$, then surely $T_r \notin F_{K'-P_C}$.

(***) Furthermore, as stated in the second bullet point above, for every non-terminal $T_u \in C$ of M-form and $\forall b \in \Gamma^u : T_u \xrightarrow{b} T_v, T_v \notin F_{K'-P_C}$.

And as we know, for every $T_v \in S_{K'-P_C} \cup Z_{K'-P_C}$, property $(B)^v_{K'-P_C}$ holds. In other words, there exists a value $g > 0$ such that regardless of strategy σ, for any $T_v \notin F_{K'-P_C}$, $Pr_{T_v}^{\sigma}[\bigcap_{q \in K'-P_C} Reach(T_q)] \leq 1 - g$.

Now let σ be an arbitrary strategy fixed for the player. Denote by w the path (in the play), where w begins at a starting non-terminal $T_i \in C$ and evolves in the following way. If the current copy o on the path w is of a L-form or a M-form non-terminal $T_j \in C$, then w follows along the unique successor of o in the play. And if the current copy o on path w is of a Q-form non-terminal $T_j \in C$ ($T_j \xrightarrow{1} T_{j'}\, T_r$ where w.l.o.g. $T_{j'} \in C$), then w follows along the child of non-terminal $T_{j'}$. If the current copy o on path w is of a non-terminal not belonging in C, then the path w terminates. Denote by $\Box C$ the event that path w is infinite, i.e., all non-terminals observed along path w are in C and path w never leaves C and never terminates. Then for any starting non-terminal $T_i \in C$:

$$\begin{aligned}
&Pr^{\sigma}_{T_i}\Big[\bigcap_{q\in K'} Reach(T_q)\Big] = Pr^{\sigma}_{T_i}\Big[\Big(\bigcap_{q\in P_C} Reach(T_q)\Big)\cap\Big(\bigcap_{q\in K'-P_C} Reach(T_q)\Big)\Big]\\
&\le Pr^{\sigma}_{T_i}\Big[\bigcap_{q\in K'-P_C} Reach(T_q)\Big] = Pr^{\sigma}_{T_i}\Big[\Big(\bigcap_{q\in K'-P_C} Reach(T_q)\Big)\cap\Box C\Big]+\\
&Pr^{\sigma}_{T_i}\Big[\Big(\bigcap_{q\in K'-P_C} Reach(T_q)\Big)\cap\neg\Box C\Big] = Pr^{\sigma}_{T_i}\Big[\Big(\bigcap_{q\in K'-P_C} Reach(T_q)\Big)\cap\neg\Box C\Big]\\
&\le \max_{T_v\notin F_{K'-P_C}} \sup_{\tau\in\Psi} Pr^{\tau}_{T_v}\Big[\bigcap_{q\in K'-P_C} Reach(T_q)\Big] \le 1-g
\end{aligned}$$

The event of reaching all target non-terminals from set $K'-P_C$ can be split into the event of reaching all targets non-terminals from set $K'-P_C$ and path w being infinite union with the event of reaching all targets non-terminals from set $K'-P_C$ and path w being finite. Moreover, $Pr^{\sigma}_{T_i}[(\bigcap_{q\in K'-P_C} Reach(T_q))\cap\Box C]=0$, due to statements (1.) and (**). The second to last inequality follows: because of statements (1.) and (**) there is zero probability from any non-terminal along path w to reach the targets from set $K'-P_C$ before event $\neg\Box C$ occurs; and also due to statement (***), once event $\neg\Box C$ occurs and path w leaves MEC, C, it terminates immediately in some non-terminal $T_v \notin C$ which also satisfies that $T_v \notin F_{K'-P_C}$. And the last inequality follows from property $(B)^v_{K'-P_C}$ for any such non-terminal $T_v \notin F_{K'-P_C}$.

And since σ was an arbitrary strategy for the player, then it follows that for any such MEC, C, in $G[Y_{K'}]$ (where $P_C \neq \emptyset$) and for any $T_i \in C$: $Pr^*_{T_i}[\bigcap_{q\in K'} Reach(T_q)] < 1$, i.e., property $(B)^i_{K'}$ holds.

Analysing MECs, C, where $P_C = \emptyset$, the argument is similar. Property (**) holds by definition of set P_C. And by property (3.), for every M-form non-terminal $T_u \in C$ and for every $b \in \Gamma^u : T_u \xrightarrow{b} T_{u'}$, $T_{u'} \in (Y_{K'} \cup S_{K'} \cup Z_{K'})$. Then because of properties (1.), (3.) and (**), it follows that for any $T_i \in C$, $\forall\sigma\in\Psi$: $Pr^{\sigma}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] \le \max_{T_{u'}\in(Y_{K'}\cup S_{K'}\cup Z_{K'})} Pr^{\sigma}_{T_{u'}}[\bigcap_{q\in K'} Reach(T_q)]$.

For non-terminals $T_{u'}$ in sets $S_{K'}$ and $Z_{K'}$, we already know by induction that property $(B)^{u'}_{K'}$ is satisfied. Moreover, from standard algorithms for MEC-decomposition, one can see that there is an ordering of the MECs in $G[Y_{K'}]$ where the bottom level (level 0) consists of MECs, C'', that in the induced subgraph $G[Y_{K'}]$ have no out-going edges from the MEC at all and for which $P_{C''} \neq K'$, and for further "levels" of MECs in the ordering the following is true: MECs or nodes that do not belong to any MEC, at level $t \ge 1$, have directed paths out of them leading to MECs (or nodes not in any

MEC) at levels $< t$. If we rank the MECs and the independent nodes (not belonging to any MEC) in $G[Y_{K'}]$, using this ordering, and use an inductive argument, it can be shown that, in the case when the above mentioned non-terminal $T_{u'}$ belongs to $Y_{K'}$ and MEC, C, has rank $t \geq 1$ in the ordering, then $T_{u'}$ belongs to a lower rank $< t$, and thus by the inductive argument, has been shown to have property $(B)^{u'}_{K'}$.

Therefore, for any non-terminal T_i in any MEC, C, in $G[Y_{K'}]$, $(B)^i_{K'}$ holds. And also by the inductive argument above for the ordering of nodes in $G[Y_{K'}]$, same holds for any non-terminal $T_i \in Y_{K'}$ not belonging to a MEC. □

Now we show that for non-terminals $T_i \in F_{K'}$, when the inner loop for subset $K' \subseteq K$ terminates, the property $(A)^i_{K'}$ is satisfied. That is:

$$\forall \varepsilon > 0,\ \exists \sigma^{\varepsilon}_{K'} \in \Psi :\ Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}\Big[\bigcap_{q \in K'} Reach(T_q)\Big] \geq 1 - \varepsilon$$

We will also show how to construct such a strategy $\sigma^{\varepsilon}_{K'}$, for a given $\varepsilon > 0$. Since we have already proved it for non-terminals in set $D_{K'}$, in the following Lemma we refer to the part of set $F_{K'}$ not containing set $D_{K'}$, i.e., to set $F_{K'} = X - S_{K'}$.

Lemma 4.6. *Every non-terminal $T_i \in F_{K'}$ satisfies property $(A)^i_{K'}$.*

Proof. Denote by $F^0_{K'}$ the initialized set of non-terminals from step II.8. Let us first observe the properties for non-terminals $T_i \in F_{K'} = X - S_{K'}$. None of them is a target non-terminal from set K', i.e., $i \notin K'$. If T_i is of L-form, then:

(L.0) if T_i belongs to a MEC, $C \subseteq F^0_{K'}$, then a non-terminal T_i generates with probability 1 as offspring some non-terminal either in set C or in set $D_{K'}$ (since L-form non-terminals in $X - S_{K'}$ do not have associated probabilistic rules to non-terminals in $S_{K'} \cup Z_{K'}$).

(L) otherwise, a non-terminal T_i generates with probability 1 as offspring some non-terminal either in set $F_{K'}$ or in set $D_{K'}$.

If T_i is of M-form, then $\forall a \in \Gamma^i :\ T_i \xrightarrow{a} T_d,\ T_d \notin D_{K'}$ and:

(M.0) if T_i belongs to a MEC, $C \subseteq F^0_{K'}$, then $\exists a^* \in \Gamma^i :\ T_i \xrightarrow{a^*} T_j,\ T_j \in C$.

(M) otherwise, $\exists a^* \in \Gamma^i :\ T_i \xrightarrow{a^*} T_j,\ T_j \in F_{K'}$.

If T_i is of Q-form (i.e., $T_i \xrightarrow{1} T_j\ T_r$), then $T_j, T_r \notin D_{K'}$. Moreover, if Q-form T_i belongs to a MEC, $C \subseteq F^0_{K'}$, then:

(Q.0) either, w.l.o.g., $T_j \in C$ and there exists some $q \in K'$ such that $T_r \in \bar{Z}_{\{q\}}$,

(Q.1) or, w.l.o.g., $T_j \in C$ and there is no $q \in K'$ such that $T_r \in \bar{Z}_{\{q\}}$.

Otherwise, if Q-form T_i does not belong to a MEC, $C \subseteq F^0_{K'}$ (i.e., $T_i \notin F^0_{K'}$), then:

(Q) w.l.o.g., $T_j \in F_{K'}$.

($\mathfrak{P}$) Let us recall that for every $q \in K'$, there is a deterministic static strategy $\sigma'_{\{q\}}$ for the player and a value $b_{\{q\}} > 0$ such that, for each non-terminal $T_r \in \bar{Z}_{\{q\}}$, $Pr^{\sigma'_{\{q\}}}_{T_r}[Reach(T_q)] \geq b_{\{q\}}$. Let $b := \min_{q\in K'}\{b_{\{q\}}\} > 0$.

Given $\varepsilon > 0$, let $\varepsilon' := (1-\sqrt{1-\varepsilon})/k$ (where $k = |K|$) and let us prove the Lemma and construct the randomized non-static strategy $\sigma^{\varepsilon}_{K'}$ inductively.

Consider the non-terminals added to set $F_{K'}$ at the initialization step II.8. during the last iteration of the inner loop. And, in particular, consider every MEC, C, added at step II.8. There is one of two reasons for why C was added to set $F^0_{K'}$.

For the first reason, suppose that $1 \leq |P_C| < l = |K'|$ and that there is a non-terminal $T_u \in C$ of M-form where $\exists b \in \Gamma^u :\ T_u \xrightarrow{b} T_{u'},\ T_{u'} \in F_{K'-P_C}$.

Consider any finite ancestor history h of height t (meaning the length of the sequence of ancestors that the history represents is t) such that h starts at a non-terminal $T_v \in C$ and all non-terminals in h belong to the MEC, C. Let o denote the non-terminal copy at the end of the ancestor history h.

If o is a copy of the non-terminal $T_u \in C$ (from above), let strategy $\sigma^{\varepsilon}_{K'}$ choose uniformly at random among actions from statement (M.0) if it is not the case that, for each $q \in P_C$, at least $d := \lceil \log_{(1-\frac{b}{k})} \varepsilon' \rceil$ copies of the Q-form non-terminals $T_j \in C \cap H_q$ have been encountered along the ancestor history h. Otherwise, $\sigma^{\varepsilon}_{K'}$ chooses deterministically action b, and therefore generates immediately a child o'' of non-terminal $T_{u'}$ (from above). In the entire subtree (subplay), rooted at o'', strategy τ is employed as if the play starts in o'', where $Pr^{\tau}_{T_{u'}}[\bigcap_{q'\in K'-P_C} Reach(T_{q'})] \geq \sqrt{1-\varepsilon}$ (exist by the induction assumption due to $T_{u'} \in F_{K'-P_C}$).

If o is of another M-form non-terminal $T_i \in C$, let $\sigma^{\varepsilon}_{K'}$ choose uniformly at random among actions from statement (M.0) and so in the next generation the single generated successor o' is of a non-terminal $T_j \in C$, where we proceed to use strategy $\sigma^{\varepsilon}_{K'}$ (that is being described).

If o is of a non-terminal $T_i \in C$ of L-form, from statement (L.0) we know that in the next generation the single generated successor o' is of some non-terminal $T_j \in C \cup D_{K'}$.

If $T_j \in D_{K'}$, then we use at o' and its subtree of descendants the randomized non-static strategy from property $(A)^j_{K'}$, that guarantees probability $\geq 1-\varepsilon$ of reaching all targets in set K', as if the play starts in o'. If $T_j \in C$, then we proceed by using the same strategy $\sigma^{\varepsilon}_{K'}$ (that is currently being described) at o'.

And if o is of a non-terminal $T_i \in C$ of Q-form ($T_i \xrightarrow{1} T_j\ T_r$), there are two cases for the two successor children o' (of non-terminal T_j) and o'' (of non-terminal T_r):

- **either** property (Q.0) is satisfied, i.e., $T_j \in C$ and $T_r \in \bar{Z}_{\{q\}}$, for some $q \in K'$. Then, in the next generation, we continue using the same strategy $\sigma^{\varepsilon}_{K'}$ (that is currently being described) at o' and for the entire subtree of play, rooted at o'', strategy $\sigma^{\varepsilon}_{K'}$ chooses uniformly at random a target non-terminal $T_q, q \in K'$, such that $T_r \in \bar{Z}_{\{q\}}$, and employs the strategy $\sigma'_{\{q\}}$ from statement ($\mathfrak{P}$) as if the play starts at o''. Note that $Pr^{\sigma^{\varepsilon}_{K'}}_{h(r,T_r)}[Reach(T_q)] \geq \frac{b}{|P_C|} \geq \frac{b}{k} > 0$, where $h(r,T_r)$ refers to the ancestor history for the right child o'' and where $|P_C| < l = |K'| \leq k = |K|$.

- **or** property (Q.1) is satisfied. Then, in the next generation, we continue using strategy $\sigma^{\varepsilon}_{K'}$ for o', whereas for o'' the strategy is irrelevant and an arbitrary one is chosen for o'' and thereafter in o'''s tree of descendants.

That concludes the description of the randomized non-static strategy $\sigma^{\varepsilon}_{K'}$ for non-terminals in MEC, C. Now we need to show that, indeed, that for any $T_i \in C$: $Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] \geq 1-\varepsilon$.

Denote by w the path (in the play) that begins at a starting non-terminal $T_i \in C$ and is defined as follows. If the current copy o on the path w is of a L-form or a M-form non-terminal $T_j \in C$, then w follows along the unique successor of o in the play. And if the current copy o on path w is of a Q-form non-terminal $T_j \in C$ ($T_j \xrightarrow{1} T_{j'}\ T_r$ where w.l.o.g. $T_{j'} \in C$), then w follows along the child of non-terminal $T_{j'}$. If the current copy o on path w is: either of a non-terminal not belonging in C; or of the non-terminal $T_{u'} \in F_{K'-P_C}$ (from above) and, for each $q \in P_C$, at least d copies of the Q-form non-terminals in set $C \cap H_q$ have already been encountered along w - then the path w terminates. Denote by $\Box C$ the event that path w (as defined) is infinite, i.e., path w never terminates, and by $\neg\Box_D C$ (respectively, $\neg\Box_{u'} C$) the event that path w is finite and terminates (according to the above definition of when it can terminate) in a copy of a non-terminal in set $D_{K'}$ (respectively, in a copy of non-terminal $T_{u'} \in F_{K'-P_C}$). Observe that under strategy $\sigma^{\varepsilon}_{K'}$ for any starting non-terminal $T_i \in C$, $P^{\sigma^{\varepsilon}_{K'}}_{T_i}[\Box C] = 0$. This is because strategy $\sigma^{\varepsilon}_{K'}$ guarantees that inside the MEC, C, there is a positive probability of reaching any non-terminal from any non-terminal. So, unless a L-form non-terminal along path w generates a non-terminal in $D_{K'}$, then the player is guaranteed to force the path to "stay" within C until, for each $q \in P_C$, at least d copies of the Q-form non-terminals in set $C \cap H_q$ have been encountered, at which point in the next copy of non-terminal T_u the player generates deterministically non-terminal $T_{u'}$. Let $p := P^{\sigma^{\varepsilon}_{K'}}_{T_i}[\neg\Box_D C]$ (note that $P^{\sigma^{\varepsilon}_{K'}}_{T_i}[\neg\Box_{u'} C] = 1 - p$).

Now under strategy $\sigma^{\varepsilon}_{K'}$ and starting at any non-terminal $T_i \in C$, with probability 1:

(i) either path w terminates in a copy o of a non-terminal in set $D_{K'}$, for which we already know that there is a strategy to reach all target non-terminals from set K' with probability $\geq 1 - \varepsilon$ (and according to $\sigma^{\varepsilon}_{K'}$ such a strategy is employed at o and its subtree of descendants). Hence, in the event of $\neg\Box_D C$, with probability $\geq 1 - \varepsilon$ all target non-terminals from set K' are contained in the generated play, i.e., $Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}[\bigcap_{q \in K'} Reach(T_q) \mid \neg\Box_D C] \geq 1 - \varepsilon$.

(ii) or, path w terminates in a copy of a non-terminal $T_{u'} \in F_{K'-P_C}$. Then, for each $q \in P_C$, with probability 1 (due to C being a MEC and due to the description of strategy $\sigma^{\varepsilon}_{K'}$) at least $d = \lceil \log_{(1-\frac{b}{k})} \varepsilon' \rceil$ copies o of the Q-form non-terminals $T_j \in C \cap H_q$ were generated along the path w. And each such copy o generates two children, o' of some non-terminal $T_{j'} \in C$ (the successor on path w) and o'' of some non-terminal $T_r \in \bar{Z}_{\{q\}}$, where o'' has independently a positive probability bounded away from zero (in fact, $\geq \frac{b}{k}$ due to the uniformly at random choice over strategies from statement ($\mathfrak{P}$), where, by Proposition 4.2, the value $b > 0$ does not depend on the history or the time when o'' is generated) to reach the respective target non-terminal T_q in a finite number of generations.

So suppose event $\neg\Box_{u'} C$ occurs and let, for each $q \in P_C$, $Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}[\Diamond_{\leq m} T_q \mid \neg\Box_{u'} C]$ denote the conditional probability, starting at a non-terminal $T_i \in C$ and under the described strategy $\sigma^{\varepsilon}_{K'}$, to reach target T_q with at most m generated copies of the Q-form non-terminals in set $C \cap H_q$ along the path w in the play, conditioned on event $\neg\Box_{u'} C$ occurring. Note that $\forall q \in P_C : Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}[\Diamond_{\leq 1} T_q \mid \neg\Box_{u'} C] \geq \frac{b}{|P_C|} \geq \frac{b}{k}$. That is, because with probability 1 under strategy $\sigma^{\varepsilon}_{K'}$, starting at a non-terminal $T_i \in C$, a copy o of a Q-form non-terminal in set $C \cap H_q$ is generated along path w and then there is a probability $\geq \frac{b}{k}$ to reach target T_q from the right child of o. It follows that for any $T_i \in C$ and any $q \in P_C$:

$$Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}[\neg\Diamond_{\leq d} T_q \mid \neg\Box_{u'} C] \leq \left(1 - \frac{b}{k}\right)^d \Leftrightarrow Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}[\Diamond_{\leq d} T_q \mid \neg\Box_{u'} C] \geq 1 - \left(1 - \frac{b}{k}\right)^d$$

Since $d \geq \log_{(1-\frac{b}{k})} \varepsilon'$, then $Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}[\Diamond_{\leq d} T_q \mid \neg\Box_{u'} C] \geq 1 - \varepsilon'$. Then for any $T_i \in C$ and any $q \in P_C$:

$$Pr_{T_i}^{\sigma_{K'}^{\varepsilon}}[Reach(T_q) \mid \neg\Box_{u'}C] \geq Pr_{T_i}^{\sigma_{K'}^{\varepsilon}}[\Diamond_{\leq d} T_q \mid \neg\Box_{u'}C] \geq 1-\varepsilon' \Leftrightarrow$$
$$Pr_{T_i}^{\sigma_{K'}^{\varepsilon}}[Reach^{\complement}(T_q) \mid \neg\Box_{u'}C] \leq \varepsilon'$$

So, by the union bound:

$$Pr_{T_i}^{\sigma_{K'}^{\varepsilon}}\Big[\bigcup_{q \in P_C} Reach^{\complement}(T_q) \,\Big|\, \neg\Box_{u'}C\Big] \leq |P_C| \cdot \varepsilon' \leq k \cdot \varepsilon' = 1-\sqrt{1-\varepsilon}$$
$$\Leftrightarrow Pr_{T_i}^{\sigma_{K'}^{\varepsilon}}\Big[\bigcap_{q \in P_C} Reach(T_q) \,\Big|\, \neg\Box_{u'}C\Big] \geq \sqrt{1-\varepsilon} \tag{4.1}$$

And in some finite number of generations, in a copy of the non-terminal T_u along path w action $b \in \Gamma^u$ is chosen deterministically, where $T_u \xrightarrow{b} T_{u'}$, $T_{u'} \in F_{K'-P_C}$. There exists $\sigma_{K'-P_C}^{1-\sqrt{1-\varepsilon}} \in \Psi$ such that $Pr_{T_{u'}}^{\sigma_{K'-P_C}^{1-\sqrt{1-\varepsilon}}}[\bigcap_{q' \in K'-P_C} Reach(T_{q'})] \geq \sqrt{1-\varepsilon}$. Then for any starting non-terminal $T_i \in C$:

$$Pr_{T_i}^{\sigma_{K'}^{\varepsilon}}\Big[\bigcap_{q' \in K'-P_C} Reach(T_{q'}) \,\Big|\, \neg\Box_{u'}C\Big] = Pr_{T_{u'}}^{\sigma_{K'-P_C}^{1-\sqrt{1-\varepsilon}}}\Big[\bigcap_{q' \in K'-P_C} Reach(T_{q'})\Big] \geq \sqrt{1-\varepsilon} \tag{4.2}$$

The equality follows from the fact that there is zero probability to reach targets from set $K'-P_C$ before path w terminates and also from the fact that strategy $\sigma_{K'}^{\varepsilon}$ utilizes strategy $\sigma_{K'-P_C}^{1-\sqrt{1-\varepsilon}}$ from the occurrence of $T_{u'}$ (when event $\neg\Box_{u'}C$ happens) as if the play starts in it.

Using (4.1) and (4.2), it follows that for any starting non-terminal $T_i \in C$:

$$Pr_{T_i}^{\sigma_{K'}^{\varepsilon}}\Big[\bigcap_{q \in K'} Reach(T_q) \,\Big|\, \neg\Box_{u'}C\Big]$$
$$= Pr_{T_i}^{\sigma_{K'}^{\varepsilon}}\Big[\bigcap_{q \in P_C} Reach(T_q) \,\Big|\, \neg\Box_{u'}C\Big] \cdot Pr_{T_i}^{\sigma_{K'}}\Big[\bigcap_{q' \in K'-P_C} Reach(T_{q'}) \,\Big|\, \neg\Box_{u'}C\Big]$$
$$= Pr_{T_i}^{\sigma_{K'}^{\varepsilon}}\Big[\bigcap_{q \in P_C} Reach(T_q) \,\Big|\, \neg\Box_{u'}C\Big] \cdot Pr_{T_{u'}}^{\sigma_{K'-P_C}^{1-\sqrt{1-\varepsilon}}}\Big[\bigcap_{q' \in K'-P_C} Reach(T_{q'})\Big] \geq (\sqrt{1-\varepsilon})^2$$
$$= 1-\varepsilon$$

And putting it all together, it follows that for any starting non-terminal $T_i \in C$:

$$Pr^{\sigma^\varepsilon_{K'}}_{T_i}\Big[\bigcap_{q\in K'} Reach(T_q)\Big] = Pr^{\sigma^\varepsilon_{K'}}_{T_i}\Big[\Big(\bigcap_{q\in K'} Reach(T_q)\Big) \cap \Box C\Big]$$

$$+ Pr^{\sigma^\varepsilon_{K'}}_{T_i}\Big[\Big(\bigcap_{q\in K'} Reach(T_q)\Big) \cap \neg\Box_D C\Big] + Pr^{\sigma^\varepsilon_{K'}}_{T_i}\Big[\Big(\bigcap_{q\in K'} Reach(T_q)\Big) \cap \neg\Box_{u'} C\Big]$$

$$= Pr^{\sigma^\varepsilon_{K'}}_{T_i}\Big[\bigcap_{q\in K'} Reach(T_q) \,\Big|\, \neg\Box_D C\Big] \cdot Pr^{\sigma^\varepsilon_{K'}}_{T_i}\Big[\neg\Box_D C\Big] +$$

$$Pr^{\sigma^\varepsilon_{K'}}_{T_i}\Big[\bigcap_{q\in K'} Reach(T_q) \,\Big|\, \neg\Box_{u'} C\Big] \cdot Pr^{\sigma^\varepsilon_{K'}}_{T_i}\Big[\neg\Box_{u'} C\Big]$$

$$\geq (1-\varepsilon)\cdot p + (1-\varepsilon)\cdot(1-p) = 1-\varepsilon$$

Now the second reason, why a MEC, C, in $G[F_{K'}]$ was added to $F^0_{K'}$ at step II.8., is if $P_C = K'$. Consider any finite ancestor history h, that starts at a non-terminal $T_v \in C$ and that all non-terminals in h belong to the MEC, C. Let o denote the non-terminal copy at the end of the ancestor history h. If o is of a L-form or Q-form non-terminal in C, let $\sigma^\varepsilon_{K'}$ behave the same way as was described before. And if o is of a M-form non-terminal $T_i \in C$, let $\sigma^\varepsilon_{K'}$ choose uniformly at random among actions from statement (M.0). So with probability 1: either a copy of a L-form non-terminal in C generates a child o' of some non-terminal in set $D_{K'}$, where $\sigma^\varepsilon_{K'}$ employs a strategy at o' and its subtree of descendants such that all targets in set K' are reached with probability $\geq 1-\varepsilon$ (such a strategy exists by the induction assumption); or, for each $q \in P_C = K'$, infinitely often copies of the Q-form non-terminals $T_j \in C \cap H_q$ are observed. In the latter case, it follows that, for each $q \in P_C = K'$, infinitely many independent copies o' of non-terminals $T_r \in \bar{Z}_{\{q\}}$ are generated, each of which has independently a positive probability bounded away from zero (again, $\geq \frac{b}{k}$ where, by Proposition 4.2, the value $b > 0$ does not depend on the history or the time when copy o' is generated) to reach the corresponding target non-terminal T_q in a finite number of generations. Hence for any $T_v \in C$, it is satisfied that $Pr^{\sigma^\varepsilon_{K'}}_{T_v}[\bigcap_{q\in K'} Reach(T_q)] \geq 1-\varepsilon$.

Therefore, for each type T_i in some MEC, $C \subseteq F^0_{K'}$, property $(A)^i_{K'}$ is satisfied.

Now consider the non-terminals T_i added to set $F_{K'}$ in step II.9. during the last iteration of the inner loop.

(i) If T_i is of L-form, then by statement (L) we know that with probability 1 non-terminal T_i in the next generation produces a single successor o' of some non-terminal $T_j \in F_{K'} \cup D_{K'}$, where by induction $(A)^j_{K'}$ holds. So using, for any given $\varepsilon > 0$, the strategy $\sigma^\varepsilon_{K'}$ from the induction assumption for each such non-terminal

T_j in the next generation as if the play starts in it, then property $(A)^i_{K'}$ is also satisfied.

(ii) If T_i is of M-form, then by statement (M), $\exists a^* \in \Gamma^i : T_i \xrightarrow{a^*} T_j, T_j \in F_{K'}$. Let $h := T_i(u, T_j)$. So, for every $\varepsilon > 0$, combining the already described strategy $\sigma^{\varepsilon}_{K'}$ for non-terminal T_j (from the induction assumption), as if the play starts in it, with the initial local choice of choosing deterministically action a^*, starting at a non-terminal T_i, we obtain an augmented strategy $\sigma^{\varepsilon}_{K'}$ for a starting non-terminal T_i such that $Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}[\bigcap_{q \in K'} Reach(T_q)] = Pr^{\sigma^{\varepsilon}_{K'}}_{h}[\bigcap_{q \in K'} Reach(T_q)] = Pr^{\sigma^{\varepsilon}_{K'}}_{T_j}[\bigcap_{q \in K'} Reach(T_q)] \geq 1 - \varepsilon$, i.e., $(A)^i_{K'}$ holds.

(iii) If T_i is of Q-form (i.e., $T_i \xrightarrow{1} T_j\ T_r$), then, by statement (Q), w.l.o.g. $T_j \in F_{K'}$, where we already know that, for every $\varepsilon > 0$, there is a strategy $\sigma^{\varepsilon}_{K'}$ such that $Pr^{\sigma^{\varepsilon}_{K'}}_{T_j}[\bigcap_{q \in K'} Reach(T_q)] \geq 1 - \varepsilon$. Let $h_l := T_i(l, T_j)$ and $h_r := T_i(r, T_r)$. Augmenting strategy $\sigma^{\varepsilon}_{K'}$ to be used from the next generation from the child of non-terminal T_j as if the play starts in it and using an arbitrary strategy from the child of non-terminal T_r, then it follows that $Pr^{\sigma^{\varepsilon}_{K'}}_{T_i}[\bigcup_{q \in K'} Reach^{\complement}(T_q)] \leq Pr^{\sigma^{\varepsilon}_{K'}}_{h_l}[\bigcup_{q \in K'} Reach^{\complement}(T_q)] \cdot Pr^{\sigma^{\varepsilon}_{K'}}_{h_r}[\bigcup_{q \in K'} Reach^{\complement}(T_q)] \leq Pr^{\sigma^{\varepsilon}_{K'}}_{T_j}[\bigcup_{q \in K'} Reach^{\complement}(T_q)] \leq \varepsilon$, resulting in property $(A)^i_{K'}$ also being satisfied. □

This completes the proof of Theorem 4.3 and the analysis of the limit-sure algorithm. The proof of Lemma 4.6 describes how to construct, for any subset $K' \subseteq K$ and any given $\varepsilon > 0$, the witness strategy $\sigma^{\varepsilon}_{K'}$ for the non-terminals in set $F_{K'}$. These non-static strategies $\sigma^{\varepsilon}_{K'}$ are described as functions that map finite ancestor histories belonging to the controller to distributions over actions available for the current non-terminal in the ancestor history, and can be described in such a form in time $(\log \frac{1}{\varepsilon})^{O(1)} \cdot 4^k \cdot |\mathcal{A}|^{O(1)}$. □

4.4 Algorithm for deciding $\overset{?}{\exists} \sigma \in \Psi : Pr^{\sigma}_{T_i}[\bigcap_{q \in K} Reach(T_q)] = 1$

In this section we present an algorithm for solving the qualitative almost-sure multi-target reachability problem for an OBMDP, $\mathcal{A}$, i.e., given a set $K \subseteq [n]$ of $k = |K|$ target non-terminals and given a starting non-terminal T_i, deciding whether there is a strategy for the player under which the probability of generating a play (tree) that contains all the target non-terminals from set K is 1. The algorithm runs in time $4^k \cdot |\mathcal{A}|^{O(1)}$, and hence is fixed-parameter tractable with respect to k.

As in the previous section, first as a preprocessing step, for each subset of targets $K' \subseteq K$, we compute the set $Z_{K'} := \{T_i \in V \mid \forall\sigma \in \Psi : Pr^{\sigma}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = 0\}$, using Algorithm 4.1. Let us also denote by AS_q, for every $q \in K$, the set of non-terminals T_j (including the target non-terminal T_q itself) for which there exists a strategy τ such that $Pr^{\tau}_{T_j}[Reach(T_q)] = 1$. Due to the equivalence between OBMDPs and BMDPs with respect to single-target reachability (see subsection 2.4.1), these sets can be computed in P-time by applying the algorithm from [ESY18, Theorem 9.3] to each target non-terminal T_q, $q \in K$.

After this preprocessing step, we apply Algorithm 4.3 to identify the non-terminals T_i for which there exists a strategy $\sigma^* \in \Psi$ such that $Pr^{\sigma^*}_{T_i}[\bigcap_{q\in K} Reach(T_q)] = 1$. Again K'_{-i} denotes the set $K' - \{i\}$. Also, to recall what notation $T_i \to T_j$ or $T_i \to \varnothing$ means, refer to the paragraph before Definition 12.

Before moving on with the proof of correctness of the algorithm, we would like to briefly and informally discuss the differences between Algorithms 4.2 and 4.3. Although the two algorithms look very similar, they differ in some crucial details.

First, the interpretation of the various sets being accumulated in the two algorithms differs, in order to correspond to the appropriate meaning in the context of almost-sure or limit-sure multi-target reachability. So even in the steps that look identical, different properties need to be proved for the accumulated sets and, hence, there are important differences in the proofs.

Furthermore, we can notice that the two algorithms differ in steps II.6. and II.8. Here is an informal intuition about this essential difference.

In Algorithm 4.2 (limit-sure multi-target reachability algorithm), step II.6. builds a MEC-decomposition of the dependency graph $G[X - S_{K'}]$, induced by the remaining non-terminals in set $X - S_{K'}$; step II.8. identifies those MECs, C, where starting at a non-terminal in C the following is observed: the branching (Q-form) non-terminals in C spawn two children each, at least one of which belongs to C, and other spawned children of the branching non-terminals in C can collectively reach a non-empty subset P_C of (or in the best case, all of) the target set K' with a positive probability (bounded away from zero); the player can choose to delay arbitrarily long the moment to select an action that "exits" C and, thus, can choose to reach the targets in set P_C with probability arbitrarily close to 1; and once the player chooses to "exit" C, it does so in a non-terminal that can limit-surely reach the rest of the targets in set $K' - P_C$.

Algorithm 4.3 Algorithm for almost-sure multi-target reachability in a given OBMDP. The output is the set $F_K = \{T_i \in V \mid \exists \sigma \in \Psi : Pr^{\sigma}_{T_i}[\bigcap_{q \in K} Reach(T_q)] = 1\}$.

I. Let $F_{\{q\}} := AS_q$ and $S_{\{q\}} := V - F_{\{q\}} - Z_{\{q\}}$, for each $q \in K$. Let $F_{\emptyset} := V$ and $S_{\emptyset} := \emptyset$.

II. For $l = 2 \ldots k$:

For every subset of target non-terminals $K' \subseteq K$ of size $|K'| = l$:

1. $D_{K'} := \{T_i \in V - Z_{K'} \mid$ one of the following holds:
   - T_i is of L-form where $i \in K'$, $T_i \not\to \varnothing$ and $\forall T_j \in V$: if $T_i \to T_j$, then $T_j \in F_{K'_{-i}}$.
   - T_i is of M-form where $i \in K'$ and $\exists a^* \in \Gamma^i : T_i \xrightarrow{a^*} T_j,\ T_j \in F_{K'_{-i}}$.
   - T_i is of Q-form ($T_i \xrightarrow{1} T_j\ T_r$) where $i \in K'$ and $\exists K_L \subseteq K'_{-i} : T_j \in F_{K_L} \wedge T_r \in F_{K'_{-i} - K_L}$.
   - T_i is of Q-form ($T_i \xrightarrow{1} T_j\ T_r$) where $\exists K_L \subset K'$ $(K_L \neq \emptyset) : T_j \in F_{K_L} \wedge T_r \in F_{K' - K_L}$.$\}$
2. Repeat until no change has occurred to $D_{K'}$:
   (a) add $T_i \notin D_{K'}$ to $D_{K'}$, if of L-form, $T_i \not\to \varnothing$ and $\forall T_j \in V$: if $T_i \to T_j$, then $T_j \in D_{K'}$.
   (b) add $T_i \notin D_{K'}$ to $D_{K'}$, if of M-form and $\exists a^* \in \Gamma^i : T_i \xrightarrow{a^*} T_j,\ T_j \in D_{K'}$.
   (c) add $T_i \notin D_{K'}$ to $D_{K'}$, if of Q-form ($T_i \xrightarrow{1} T_j\ T_r$) and $T_j \in D_{K'} \vee T_r \in D_{K'}$.
3. Let $X := V - (D_{K'} \cup Z_{K'})$.
4. Initialize $S_{K'} := \{T_i \in X \mid$ either $i \in K'$, or T_i is of L-form and $T_i \to \varnothing \vee T_i \to T_j,\ T_j \in Z_{K'}\} \cup \bigcup_{\emptyset \subset K'' \subset K'} (X \cap S_{K''})$.
5. Repeat until no change has occurred to $S_{K'}$:
   (a) add $T_i \in X - S_{K'}$ to $S_{K'}$, if of L-form and $T_i \to T_j,\ T_j \in S_{K'} \cup Z_{K'}$.
   (b) add $T_i \in X - S_{K'}$ to $S_{K'}$, if of M-form and $\forall a \in \Gamma^i :\ T_i \xrightarrow{a} T_j,\ T_j \in S_{K'} \cup Z_{K'}$.
   (c) add $T_i \in X - S_{K'}$ to $S_{K'}$, if of Q-form ($T_i \xrightarrow{1} T_j\ T_r$) and $T_j \in S_{K'} \cup Z_{K'} \wedge T_r \in S_{K'} \cup Z_{K'}$.
6. $\mathcal{C} \leftarrow$ SCC decomposition of $G[X - S_{K'}]$.
7. For every $q \in K'$, let $H_q := \{T_i \in X - S_{K'} \mid T_i$ is of Q-form ($T_i \xrightarrow{1} T_j\ T_r$) and $((T_j \in X - S_{K'} \wedge T_r \in \bar{Z}_{\{q\}}) \vee (T_j \in \bar{Z}_{\{q\}} \wedge T_r \in X - S_{K'}))\}$.
8. Let $F_{K'} := \bigcup \{\cup_{q \in K'} (H_q \cap C) \mid C \in \mathcal{C}$ s.t. $\forall q' \in K' : H_{q'} \cap C \neq \emptyset\}$.
9. Repeat until no change has occurred to $F_{K'}$:
   (a) add $T_i \in X - (S_{K'} \cup F_{K'})$ to $F_{K'}$, if of L-form and $T_i \to T_j,\ T_j \in F_{K'} \cup D_{K'}$.
   (b) add $T_i \in X - (S_{K'} \cup F_{K'})$ to $F_{K'}$, if of M-form and $\exists a^* \in \Gamma^i : T_i \xrightarrow{a^*} T_j,\ T_j \in F_{K'}$.
   (c) add $T_i \in X - (S_{K'} \cup F_{K'})$ to $F_{K'}$, if of Q-form ($T_i \xrightarrow{1} T_j\ T_r$) and $T_j \in F_{K'} \vee T_r \in F_{K'}$.
10. If $X \neq S_{K'} \cup F_{K'}$, let $S_{K'} := X - F_{K'}$ and go to step 5.
11. Else, i.e., if $X = S_{K'} \cup F_{K'}$, let $F_{K'} := F_{K'} \cup D_{K'}$.

III. **Output** F_K.

On the other hand, in Algorithm 4.3 (almost-sure multi-target reachability algorithm), step II.6. builds a SCC-decomposition of the dependency graph $G[X - S_{K'}]$, induced by the remaining non-terminals in set $X - S_{K'}$; step II.8. identifies those branching (Q-form) non-terminals that belong to SCCs, C, where the following is true for each such C: the Q-form non-terminals in C (that have been identified in step II.8.) spawn two children each, at least one of which belongs to C, and the other spawned children of these same branching non-terminals can collectively reach all the targets in set K' with a positive probability (bounded away from zero).

Now let us continue with the proof of correctness of Algorithm 4.3 and the theorem behind it.

Theorem 4.7. *Algorithm 4.3 computes, given an OBMDP, $\mathcal{A}$, and a set $K \subseteq [n]$ of $k = |K|$ target non-terminals, for each subset $K' \subseteq K$, the set of non-terminals $F_{K'} := \{T_i \in V \mid \exists \sigma \in \Psi:\ Pr^{\sigma}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = 1\}$. The algorithm runs in time $4^k \cdot |\mathcal{A}|^{O(1)}$. Moreover, for each $K' \subseteq K$, the algorithm can also be augmented to compute a randomized non-static strategy $\sigma^*_{K'}$ such that $Pr^{\sigma^*_{K'}}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = 1$ for all non-terminals $T_i \in F_{K'}$.*

Proof. We will refer to the loop executing steps II.5. through II.10. for a specific subset $K' \subseteq K$ as the "inner" loop and the iteration through all subsets of K as the "outer" loop. Clearly the inner loop terminates, due to step II.10. always adding at least one non-terminal to set $S_{K'}$ and step II.11. eventually executing. The running time of the algorithm follows from the facts that the outer loop executes for 2^k iterations and inside each iteration of the outer loop, steps II.1. and II.4. require time at most $2^k \cdot |\mathcal{A}|^{O(1)}$ and the inner loop executes for at most $|V|$ iterations, where during each inner loop iteration the steps in it execute in time at most $|\mathcal{A}|^{O(1)}$.

For the proof of correctness, we show that for every subset of target non-terminals $K' \subseteq K$, $F_{K'}$ (from the decomposition $V = F_{K'} \cup S_{K'} \cup Z_{K'}$) is the set of non-terminals T_i for which the following property holds:

$$(A)^i_{K'}\text{: } \exists \sigma_{K'} \in \Psi \text{ such that } Pr^{\sigma_{K'}}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = 1.$$

Otherwise, if $T_i \in S_{K'}$, then we show that the following property holds:

$(B)^i_{K'}$: $\forall \sigma \in \Psi:\ Pr^{\sigma}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] < 1 \Leftrightarrow Pr^{\sigma}_{T_i}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] > 0$, i.e., the probability of generating a play that contains at least one copy for each of the T_q ($q \in K'$) target non-terminals, is < 1.

Clearly, for non-terminals $T_i \in Z_{K'}$, property $(B)^i_{K'}$ holds because, by Proposition 4.2, $\forall \sigma \in \Psi: \; Pr^{\sigma}_{T_i}[\bigcap_{q \in K'} Reach(T_q)] = 0 < 1$. Finally, the answer for the full set of targets is $F := F_K$.

As in the proof from the previous section, we base this proof on an induction on the size of subset K', i.e. on the time of computing sets $S_{K'}$ and $F_{K'}$ for $K' \subseteq K$. And in the process, for each subset $K' \subseteq K$ of target non-terminals, we construct a randomized *non-static* strategy $\sigma_{K'}$ for the player that ensures $Pr^{\sigma_{K'}}_{T_i}[\bigcap_{q \in K'} Reach(T_q)] = 1$ for each non-terminal $T_i \in F_{K'}$. In the end, $\sigma := \sigma_K$ is the strategy that guarantees almost-sure reachability of all given targets in the same play.

To begin with, observe that clearly for any subset of target non-terminals, $K' := \{q\} \subseteq K$, of size $l = 1$, each non-terminal $T_i \in F_{\{q\}}$ (respectively, $T_i \in V - F_{\{q\}}$) satisfies property $(A)^i_{\{q\}}$ (respectively, $(B)^i_{\{q\}}$), due to step I. and the definition of the $AS_q, q \in K$ sets. Hence, for each such subset $\{q\} \subseteq K$, there is a strategy $\sigma_{\{q\}}$ such that $\forall T_i \in F_{\{q\}} : \; Pr^{\sigma_{\{q\}}}_{T_i}[Reach(T_q)] = 1$. Moreover, by [ESY18, Theorem 9.4] this strategy $\sigma_{\{q\}}$ is non-static and deterministic. Analysing subset K' of target non-terminals of size l as part of step II., assume that, for every $K'' \subset K'$ of size $\leq l-1$, sets $S_{K''}$ and $F_{K''}$ have already been computed, and for each non-terminal T_j belonging to set $F_{K''}$ (respectively, set $S_{K''}$) property $(A)^j_{K''}$ (respectively, $(B)^j_{K''}$) holds. That is, by induction assumption, for each $K'' \subset K'$, there is a randomized non-static strategy $\sigma_{K''}$ such that for any $T_j \in F_{K''}$: $Pr^{\sigma_{K''}}_{T_j}[\bigcap_{q \in K''} Reach(T_q)] = 1$, and also for any $T_j \in S_{K''}$: $\forall \sigma \in \Psi$, $Pr^{\sigma}_{T_j}[\bigcap_{q \in K''} Reach(T_q)] < 1$. We now need to show that at end of the inner loop analysis of subset K', property $(A)^i_{K'}$ (respectively, $(B)^i_{K'}$) holds for every non-terminal $T_i \in F_{K'}$ (respectively, $T_i \in S_{K'}$).

First we show that property $(A)^i_{K'}$ holds for each non-terminal T_i belonging to set $D_{K'}$ ($\subseteq F_{K'}$), precomputed prior to the execution of the inner loop for subset K'.

Lemma 4.8. *Every non-terminal $T_i \in D_{K'}$ satisfies property $(A)^i_{K'}$.*

Proof. The lemma is proved via a nested induction based on the time of a non-terminal being added to set $D_{K'}$. Consider the base case where $T_i \in D_{K'}$ is a non-terminal, added at the initialization step II.1.

(i) Suppose T_i is of L-form where $i \in K'$ and for all associated rules a child is generated that is of a non-terminal $T_j \in F_{K'_{-i}}$, where property $(A)^j_{K'_{-i}}$ holds. Then using the witness strategy from property $(A)^j_{K'_{-i}}$, that almost-surely reaches all remaining targets from set K'_{-i}, for all such non-terminals T_j in the next generation as if the play starts in it and, since the starting target non-terminal T_i is already reached, clearly property $(A)^i_{K'}$ holds.

(ii) Suppose T_i is of M-form where $i \in K'$ and $\exists a^* \in \Gamma^i$ such that $T_i \xrightarrow{a^*} T_j$, $T_j \in F_{K'_{-i}}$, where property $(A)^j_{K'_{-i}}$ holds by induction. Let $h := T_i(u, T_j)$. Then, by combining the witness strategy $\sigma_{K'_{-i}}$ from the induction assumption for non-terminal T_j, as if the play starts in it, with the initial local choice of choosing deterministically action a^* starting at a non-terminal T_i, we obtain a combined strategy $\sigma_{K'}$ such that starting at a (target) non-terminal T_i, $Pr^{\sigma_{K'}}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = Pr^{\sigma_{K'}}_{T_i}[\bigcap_{q\in K'_{-i}} Reach(T_q) \mid Reach(T_i)] \cdot Pr^{\sigma_{K'}}_{T_i}[Reach(T_i)] = Pr^{\sigma_{K'}}_{T_i}[\bigcap_{q\in K'_{-i}} Reach(T_q)] = Pr^{\sigma_{K'}}_{h}[\bigcap_{q\in K'_{-i}} Reach(T_q)] = Pr^{\sigma_{K'_{-i}}}_{T_j}[\bigcap_{q\in K'_{-i}} Reach(T_q)] = 1$.

(iii) Suppose T_i is of Q-form ($T_i \xrightarrow{1} T_j\ T_r$) where $i \in K'$ and there exists a split of the rest of the target non-terminals, implied by $K_L \subseteq K'_{-i}$ and $K'_{-i} - K_L$, such that $T_j \in F_{K_L} \wedge T_r \in F_{K'_{-i}-K_L}$. Let $h_l := T_i(l, T_j)$ and $h_r := T_i(r, T_r)$. By combining the two witness strategies σ_{K_L} and $\sigma_{K'_{-i}-K_L}$ from the induction assumption for non-terminals T_j and T_r, respectively, to be used from the next generation as if the play starts in it, and the fact that target T_i is reached (since T_i is the starting non-terminal), it follows that there exists a strategy $\sigma_{K'} \in \Psi$ such that $Pr^{\sigma_{K'}}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = Pr^{\sigma_{K'}}_{T_i}[\bigcap_{q\in K'_{-i}} Reach(T_q)] \geq Pr^{\sigma_{K'}}_{h_l}[\bigcap_{q\in K_L} Reach(T_q)] \cdot Pr^{\sigma_{K'}}_{h_r}[\bigcap_{q\in K'_{-i}-K_L} Reach(T_q)] = Pr^{\sigma_{K_L}}_{T_j}[\bigcap_{q\in K_L} Reach(T_q)] \cdot Pr^{\sigma_{K'_{-i}-K_L}}_{T_r}[\bigcap_{q\in K'_{-i}-K_L} Reach(T_q)] = 1$.

(iv) Suppose T_i is of Q-form ($T_i \xrightarrow{1} T_j\ T_r$) where there exists a proper split of the target non-terminals from set K', implied by $K_L \subset K'$ (where $K_L \neq \emptyset$) and $K' - K_L$, such that $T_j \in F_{K_L} \wedge T_r \in F_{K'-K_L}$. Combining the two witness strategies σ_{K_L} and $\sigma_{K'-K_L}$ from the induction assumption for non-terminals T_j, T_r in the same way as in (iii), it follows that there exists a strategy $\sigma_{K'} \in \Psi$ such that property $(A)^i_{K'}$ holds.

Now consider non-terminals T_i added to set $D_{K'}$ at step II.2., i.e., the inductive step. If non-terminal T_i is of L-form, then all rules, associated with it, generate children of non-terminals T_j already in set $D_{K'}$, for which $(A)^j_{K'}$ holds by the (nested) induction. Hence, $(A)^i_{K'}$ clearly also holds for the same reason as in (i) above.

If non-terminal T_i is of M-form, then $\exists a^* \in \Gamma^i : T_i \xrightarrow{a^*} T_j, T_j \in D_{K'}$. Again let $h := T_i(u, T_j)$. By combining the witness strategy $\sigma_{K'}$ for non-terminal T_j (from the nested induction assumption), as if the play starts in it, with the initial local choice of choosing deterministically action a^* starting at a non-terminal T_i, we obtain an augmented strategy $\sigma_{K'}$ for a starting non-terminal T_i such that $Pr_{T_i}^{\sigma_{K'}}[\bigcap_{q\in K'} Reach(T_q)] = Pr_h^{\sigma_{K'}}[\bigcap_{q\in K'} Reach(T_q)] = Pr_{T_j}^{\sigma_{K'}}[\bigcap_{q\in K'} Reach(T_q)] = 1$.

If T_i is of Q-form ($T_i \xrightarrow{1} T_j\ T_r$), then either $T_j \in D_{K'}$ or $T_r \in D_{K'}$, i.e., $\exists \sigma_{K'} \in \Psi$ such that $Pr_{T_y}^{\sigma_{K'}}[\bigcap_{q\in K'} Reach(T_q)] = 1 \Leftrightarrow Pr_{T_y}^{\sigma_{K'}}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] = 0$, where $y \in \{j,r\}$. Let $h_y := T_i(x, T_y)$ and $h_{\bar{y}} := T_i(\bar{x}, T_{\bar{y}})$, where $\bar{y} \in \{j,r\} - \{y\}$, $x \in \{l,r\}$ and $\bar{x} \in \{l,r\} - \{x\}$. By augmenting this $\sigma_{K'}$ to be used from the next generation from the child of non-terminal T_y, as if the play starts in it, and using an arbitrary strategy from the child of non-terminal $T_{\bar{y}}$, it follows that $Pr_{T_i}^{\sigma_{K'}}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \leq Pr_{h_y}^{\sigma_{K'}}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \cdot Pr_{h_{\bar{y}}}^{\sigma_{K'}}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \leq Pr_{T_y}^{\sigma_{K'}}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] = 0$, i.e., property $(A)_{K'}^i$ is satisfied. □

Next we show that if $T_i \in S_{K'}$, then property $(B)_{K'}^i$ holds.

Lemma 4.9. *Every non-terminal $T_i \in S_{K'}$ satisfies property $(B)_{K'}^i$.*

Proof. This can be done again via another (nested) induction, based on the time a non-terminal is added to set $S_{K'}$. That is, assuming all non-terminals T_j, added already to set $S_{K'}$ in previous steps and iterations of the inner loop, satisfy property $(B)_{K'}^j$, then we show that for a new addition T_i to set $S_{K'}$, property $(B)_{K'}^i$ is also satisfied.

Consider the initialized set $S_{K'}$ of non-terminals T_i constructed at step II.4.

If T_i is of L-form, where $T_i \to \varnothing \vee T_i \to T_j$, $T_j \in Z_{K'}$, then with a positive probability non-terminal T_i immediately either does not generate a child at all or generates a child of non-terminal $T_j \in Z_{K'}$, for which we already know that $(B)_{K'}^j$ holds. Clearly, this also results in $(B)_{K'}^i$ being satisfied.

If, for some subset $K'' \subset K'$, $T_i \in S_{K''}$, then $\forall \sigma \in \Psi:\ Pr_{T_i}^{\sigma}[\bigcup_{q\in K''} Reach^{\complement}(T_q)] > 0$ (i.e., property $(B)_{K''}^i$). But, $\forall \sigma \in \Psi:\ Pr_{T_i}^{\sigma}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \geq Pr_{T_i}^{\sigma}[\bigcup_{q\in K''} Reach^{\complement}(T_q)] > 0$, so property $(B)_{K'}^i$ also holds. Note that if, for some subset $K'' \subset K'$, $T_i \in Z_{K''}$, then $T_i \in Z_{K'}$ and so already $T_i \notin X$.

And if T_i is a target non-terminal in set K', then due to not being added to set $D_{K'}$ in step II.1. it follows that: (1) if of L-form, it generates with a positive probability a child of a non-terminal $T_j \in S_{K'_{-i}} \cup Z_{K'_{-i}}$, for which $(B)_{K'_{-i}}^j$ holds; (2) if of M-form, irrespective of the strategy it generates a child of a non-terminal $T_j \in S_{K'_{-i}} \cup Z_{K'_{-i}}$, for which again $(B)_{K'_{-i}}^j$ holds; (3) and if of Q-form, it generates two children of non-terminals T_j, T_r, for which no matter how we split the rest of the target non-terminals in set K'_{-i} (into subsets $K_L \subseteq K'_{-i}$ and $K'_{-i} - K_L$), either $(B)_{K_L}^j$ holds or $(B)_{K'_{-i}-K_L}^r$ holds. In other words, a target non-terminal T_i in the initial set $S_{K'}$ has no strategy to ensure that the rest of the target non-terminals are reached with probability 1 (the reasoning behind this last statement is the same as the arguments in (i) - (iii) below, since for a starting (target) non-terminal T_i: $\forall\sigma \in \Psi : \ Pr^{\sigma}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = Pr^{\sigma}_{T_i}[\bigcap_{q\in K'_{-i}} Reach(T_q)]$).

Observe that by the end of step II.4. all target non-terminals $T_q, q \in K'$ belong either to set $D_{K'}$ or set $S_{K'}$. Now consider a non-terminal T_i added to set $S_{K'}$ in step II.5. during some iteration of the inner loop.

(i) Suppose T_i is of L-form. Then $T_i \to T_j$, $T_j \in S_{K'} \cup Z_{K'}$, where property $(B)^j_{K'}$ holds. So regardless of the strategy σ for the player, there is a positive probability to generate a child of the above non-terminal T_j, where $Pr^{\sigma}_{T_j}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] > 0$. Let $h := T_i(u, T_j)$. But note that, $\forall\sigma \in \Psi$: $Pr^{\sigma}_{T_i}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \geq p_{i,j} \cdot Pr^{\sigma}_{h}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] > 0$ if and only if $\forall\sigma \in \Psi : \ p_{i,j} \cdot Pr^{\sigma}_{T_j}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] > 0$, where $p_{i,j} > 0$ is the probability of the rule $T_i \xrightarrow{p_{i,j}} T_j$. And since the latter part of the statement holds, then the former (i.e., property $(B)^i_{K'}$) is satisfied.

(ii) Suppose T_i is of M-form. Then $\forall a \in \Gamma^i : \ T_i \xrightarrow{a} T_j$, $T_j \in S_{K'} \cup Z_{K'}$. So irrelevant of strategy σ, starting in a non-terminal T_i, the next generation surely consists of some non-terminal T_j such that $\forall\sigma \in \Psi : \ Pr^{\sigma}_{T_j}[\bigcap_{q\in K'} Reach(T_q)] < 1$. Clearly, $\forall\sigma \in \Psi : \ Pr^{\sigma}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] \leq \max_{\{T_j \in S_{K'} \cup Z_{K'} | T_i \to T_j\}} Pr^{\sigma}_{T_i(u,T_j)}[\bigcap_{q\in K'} Reach(T_q)] < 1$ (i.e., property $(B)^i_{K'}$) if and only if $\forall\sigma \in \Psi : \max_{\{T_j \in S_{K'} \cup Z_{K'} | T_i \to T_j\}} Pr^{\sigma}_{T_j}[\bigcap_{q\in K'} Reach(T_q)] < 1$, where the latter is satisfied.

(iii) Suppose T_i is of Q-form (i.e., $T_i \xrightarrow{1} T_j\ T_r$). Then $T_j, T_r \in S_{K'} \cup Z_{K'}$, i.e., both $(B)^j_{K'}$ and $(B)^r_{K'}$ are satisfied. We know that:

1) Neither of the children can single-handedly reach all target non-terminals from set K' with probability 1. That is, $\forall\sigma \in \Psi$, $Pr^{\sigma}_{T_j}[\bigcap_{q\in K'} Reach(T_q)] < 1$ and $Pr^{\sigma}_{T_r}[\bigcap_{q\in K'} Reach(T_q)] < 1$.

2) Moreover, since T_i was not added to set $D_{K'}$ in step II.1., then $\forall K_L \subset K'$ (where $K_L \neq \emptyset$) either $(B)^j_{K_L}$ holds (i.e., $T_j \notin F_{K_L}$) or $(B)^r_{K'-K_L}$ holds (i.e., $T_r \notin F_{K'-K_L}$), i.e., either $\forall\sigma \in \Psi : \ Pr^{\sigma}_{T_j}[\bigcap_{q\in K_L} Reach(T_q)] < 1$ or $\forall\sigma \in \Psi : Pr^{\sigma}_{T_r}[\bigcap_{q\in K'-K_L} Reach(T_q)] < 1$.

Let $h_l := T_i(l, T_j)$ and $h_r := T_i(r, T_r)$. Notice that for any strategy $\sigma \in \Psi$ and for any $q' \in K'$, $Pr^{\sigma}_{T_i}[\bigcup_{q\in K'} Reach^{\complement}(T_q)] \geq Pr^{\sigma}_{T_i}[Reach^{\complement}(T_{q'})] = Pr^{\sigma}_{h_l}[Reach^{\complement}(T_{q'})] \cdot Pr^{\sigma}_{h_r}[Reach^{\complement}(T_{q'})]$.

We claim that $\forall \sigma \in \Psi : \bigvee_{q \in K'} Pr^{\sigma}_{T_j}[Reach^{\complement}(T_q)] \cdot Pr^{\sigma}_{T_r}[Reach^{\complement}(T_q)] > 0$. But for any $q \in K'$ and for any $\sigma \in \Psi$ one can easily construct $\sigma' \in \Psi$ such that $Pr^{\sigma}_{T_j}[Reach^{\complement}(T_q)] = Pr^{\sigma'}_{h_l}[Reach^{\complement}(T_q)]$ and similarly for non-terminal T_r. So it follows from the claim that $\forall \sigma \in \Psi : \bigvee_{q \in K'} Pr^{\sigma}_{h_l}[Reach^{\complement}(T_q)] \cdot Pr^{\sigma}_{h_r}[Reach^{\complement}(T_q)] > 0$ and, so that $\forall \sigma \in \Psi : Pr^{\sigma}_{T_i}[\bigcup_{q \in K'} Reach^{\complement}(T_q)] > 0 \Leftrightarrow Pr^{\sigma}_{T_i}[\bigcap_{q \in K'} Reach(T_q)] < 1$.

Suppose the opposite, i.e., assume $(\mathcal{P})$ that $\exists \sigma' \in \Psi : \bigwedge_{q \in K'} Pr^{\sigma'}_{T_j}[Reach^{\complement}(T_q)] \cdot Pr^{\sigma'}_{T_r}[Reach^{\complement}(T_q)] = 0$. Now for any $q \in K'$, by statement 2) above, we know that $T_j \notin F_{\{q\}} \vee T_r \notin F_{K'_{-q}}$ and $T_j \notin F_{K'_{-q}} \vee T_r \notin F_{\{q\}}$. First, suppose that in fact for some $q' \in K'$ it is the case that $T_j \notin F_{\{q'\}} \wedge T_r \notin F_{\{q'\}}$ (i.e., $T_j \in S_{\{q'\}} \cup Z_{\{q'\}} \wedge T_r \in S_{\{q'\}} \cup Z_{\{q'\}}$). That is, $\forall \sigma \in \Psi : Pr^{\sigma}_{T_j}[Reach^{\complement}(T_{q'})] > 0$ and $Pr^{\sigma}_{T_r}[Reach^{\complement}(T_{q'})] > 0$, where our claims follows directly (hence, contradiction to $(\mathcal{P})$). Second, suppose that for some $q' \in K'$ it is the case that $T_j \notin F_{K'_{-q'}} \wedge T_r \notin F_{K'_{-q'}}$ (i.e., $T_j \in S_{K'_{-q'}} \cup Z_{K'_{-q'}} \wedge T_r \in S_{K'_{-q'}} \cup Z_{K'_{-q'}}$). But then T_i would have been added to set $S_{K'_{-q'}}$ at step II.5.(c) when constructing the answer for subset of targets $K'_{-q'}$. However, we already know that $T_i \in \bigcap_{K'' \subset K'} F_{K''}$ (follows from steps II.3 and II.4. that $T_i \notin \bigcup_{K'' \subset K'} (S_{K''} \cup Z_{K''})$). Hence, again a contradiction.

Therefore, it follows that for every $q \in K'$, either $T_j \notin F_{\{q\}} \wedge T_j \notin F_{K'_{-q}}$ or $T_r \notin F_{\{q\}} \wedge T_r \notin F_{K'_{-q}}$. And in particular, the essential part is that $\forall q \in K'$, either $T_j \notin F_{\{q\}}$ or $T_r \notin F_{\{q\}}$. That is, for every $q \in K'$, either $\forall \sigma \in \Psi : Pr^{\sigma}_{T_j}[Reach^{\complement}(T_q)] > 0$, or $\forall \sigma \in \Psi : Pr^{\sigma}_{T_r}[Reach^{\complement}(T_q)] > 0$. But then, combined with assumption $(\mathcal{P})$, it actually follows that there exists a subset $K'' \subseteq K'$ such that $\exists \sigma' \in \Psi : \bigwedge_{q \in K''} Pr^{\sigma'}_{T_r}[Reach^{\complement}(T_q)] = 0 \wedge \bigwedge_{q \in K' - K''} Pr^{\sigma'}_{T_j}[Reach^{\complement}(T_q)] = 0$. And by Proposition 2.2(1.), it follows that there exists a $\sigma' \in \Psi$ such that $Pr^{\sigma'}_{T_r}[\bigcap_{q \in K''} Reach(T_q)] = 1 \wedge Pr^{\sigma'}_{T_j}[\bigcap_{q \in K' - K''} Reach(T_q)] = 1$, i.e., $T_j \in F_{K' - K''} \wedge T_r \in F_{K''}$, contradicting the known facts 1) and 2). Hence, assumption $(\mathcal{P})$ is wrong and our claim is satisfied.

Now consider any non-terminal T_i that is added to set $S_{K'}$ in step II.10. at some iteration of the inner loop (i.e., $T_i \in Y_{K'} := X - (S_{K'} \cup F_{K'}) \subseteq \bar{Z}_{K'}$). Since non-terminal T_i has not been previously added to sets $D_{K'}$, $S_{K'}$ or $F_{K'}$, then all of the following hold:

(1.) $i \notin K'$;

(2.) if T_i is of L-form, then a non-terminal T_i generates with probability 1 a non-terminal which belongs to $Y_{K'}$ (otherwise T_i would have been added to sets $S_{K'}$ or $F_{K'}$ in step II.4, II.5. or II.9., respectively);

(3.) if T_i is of M-form, then $\forall a \in \Gamma^i: \; T_i \xrightarrow{a} T_d, \; T_d \notin F_{K'} \cup D_{K'}$ (otherwise T_i would have been added to sets $D_{K'}$ or $F_{K'}$ in step II.2. or step II.9., respectively), and $\exists a' \in \Gamma^i: \; T_i \xrightarrow{a'} T_j, \; T_j \notin S_{K'} \cup Z_{K'}$, i.e., $T_j \in Y_{K'}$ (otherwise T_i would have been added to set $S_{K'}$ in step II.5.); and

(4.) if T_i is of Q-form ($T_i \xrightarrow{1} T_j \; T_r$), then w.l.o.g. $T_j \in Y_{K'}$ and $T_r \in (Y_{K'} \cup S_{K'} \cup Z_{K'})$ (as T_i has not been added to the other sets in steps II.2., II.5. or II.9.).

Due to the statements (2.) - (4.) above, notice that the dependency graph G does not contain outgoing edges from set $Y_{K'}$ to sets $D_{K'}$ and $F_{K'}$. So any SCC in subgraph $G[X - S_{K'}]$, that contains a node from set $Y_{K'}$, is in fact entirely contained in subgraph $G[Y_{K'}]$.

Furthermore, one of the following is the reason for a Q-form non-terminal $T_i \in Y_{K'}$ ($T_i \xrightarrow{1} T_j \; T_r$) not having been added to set $F_{K'}$ at the initialization step II.8.:

(4.1.) either T_i does not belong to any of the sets $H_q, q \in K'$. So, from step II.7., $T_r \in Z_{\{q\}}$ for every $q \in K'$ (recall from property (4.) that w.l.o.g. $T_j \in Y_{K'} \subseteq \bar{Z}_{K'} \subseteq \bar{Z}_{\{q\}}, \; \forall q \in K'$),

(4.2.) or T_i does belong to some set $H_{q'}, q' \in K'$, but if T_i belongs to a strongly connected component C' in $G[Y_{K'}]$, then $\exists q'' \in K'$ such that $H_{q''} \cap C' = \emptyset$.

We can treat the Q-form non-terminals with property (4.1.) as if they have only one child (namely the child of non-terminal T_j), since the other child (of non-terminal T_r) does not contribute to reaching, even with a positive probability, any of the target non-terminals from set K'.

We need to show that for every non-terminal $T_i \in Y_{K'}$ property $(B)^i_{K'}$ holds, i.e., $\forall \sigma \in \Psi: \; Pr^{\sigma}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] < 1$.

From standard algorithms about SCC-decomposition, it is known that there is an ordering of the SCCs in $G[Y_{K'}]$, where the bottom level in this ordering (level 0) consists of strongly connected components that in the induced subgraph $G[Y_{K'}]$ have no edges leaving the SCC at all, and for further levels in the ordering of SCCs the following is true: SCCs or nodes not in any SCC, at level $t \geq 1$, have directed paths out of them leading to SCCs or nodes not in any SCC, at levels $< t$. We rank the SCCs and the independent nodes (not belonging to any SCC) in $G[Y_{K'}]$ according to this ordering, denoting by $Y^t_{K'}$, $t \geq 0$, the nodes (non-terminals) at levels up to and including t, and use the following induction based on the level:

– For the base case: for any SCC, C, at level 0 (i.e., $C \subseteq Y^0_{K'}$), clearly for any non-terminal $T_i \in C$, $\exists \sigma \in \Psi : Pr^{\sigma}_{T_i}[\bigcap_{q \in K'} Reach(T_q)] = 1$ if and only if $H_q \cap C \neq \emptyset$, $\forall q \in K'$. But, by property (4.2), there is no such component C in $G[Y_{K'}]$ that contains a Q-form non-terminal from each of the sets H_q, $q \in K'$.

– As for the inductive step, assume that for some $t \geq 1$ for any $T_v \in Y^{t-1}_{K'}$, $\forall \sigma \in \Psi : Pr^{\sigma}_{T_v}[\bigcap_{q \in K'} Reach(T_q)] < 1$, i.e., $(B)^v_{K'}$ is satisfied. Let σ be an arbitrary strategy fixed for the player. For a SCC, C', at level $t \geq 1$, let w denote the path (in the play), where w begins at a starting non-terminal $T_i \in C'$ and evolves in the following way. If the current copy o on the path w is of a L-form or a M-form non-terminal $T_j \in C'$, then w follows along the unique successor of o in the play. And if the current copy o on the path w is of a Q-form non-terminal $T_j \in C'$ ($T_j \xrightarrow{1} T_{j'}\, T_r$, where w.l.o.g. $T_{j'} \in C'$), then w follows along the child of non-terminal $T_{j'}$. (Note that $T_r \notin C'$, since we already know from (4.) that $T_{j'} \in Y_{K'} \subseteq \bar{Z}_{K'} \subseteq \bar{Z}_{\{q\}}$, $\forall q \in K'$, and so if $T_r \in C' \subseteq Y_{K'}$ then property (4.2.) will be contradicted.) If the current copy o on the path w is of a non-terminal not belonging in C', then the path w terminates. Denote by $\Box C'$ the event that path w is infinite, i.e., all non-terminals observed along path w are in C' and path w never leaves C' and never terminates. Then for any starting non-terminal $T_i \in C'$:

$$Pr^{\sigma}_{T_i}\Big[\bigcap_{q \in K'} Reach(T_q)\Big] = Pr^{\sigma}_{T_i}\Big[\Big(\bigcap_{q \in K'} Reach(T_q)\Big) \cap \Box C'\Big]$$
$$+ Pr^{\sigma}_{T_i}\Big[\Big(\bigcap_{q \in K'} Reach(T_q)\Big) \cap \neg\Box C'\Big] = Pr^{\sigma}_{T_i}\Big[\Big(\bigcap_{q \in K'} Reach(T_q)\Big) \cap \neg\Box C'\Big]$$

Observe that $Pr^{\sigma}_{T_i}[(\bigcap_{q \in K'} Reach(T_q)) \cap \Box C'] = 0$, due to statements (1.) and (4.2.).

By property (3.) and also due to the ranking of SCCs and nodes in $G[Y_{K'}]$, if path w terminates, then it does in a non-terminal $T_v \in S_{K'} \cup Z_{K'} \cup Y^{t-1}_{K'}$. Also due to properties (1.) - (4.) and (4.2.), in the case of event $\neg\Box C'$ occurring, all the targets in set K' are reached with probability 1, starting in $T_i \in C'$, if and only if they are all reached with probability 1, starting from such a non-terminal $T_v \in S_{K'} \cup Z_{K'} \cup Y^{t-1}_{K'}$. To see this, note that for any of the Q-form non-terminals $T_j \in C'$ ($T_j \xrightarrow{1} T_{j'}\, T_r$, where w.l.o.g. $T_{j'} \in C'$), $T_r \notin F_{\{q\}}$ for any $q \in K'$ (otherwise, if $T_r \in F_{\{q'\}}$ for some $q' \in K'$, then since T_j was not added to set $D_{K'}$ at step II.1., it follows that $T_{j'} \notin F_{K'_{-q'}}$, i.e., $T_{j'} \in S_{K'_{-q'}} \cup Z_{K'_{-q'}}$, and hence by step (4.) of algorithm already $T_{j'} \in S_{K'} \cup Z_{K'}$, which contradicts that $T_{j'} \in C' \subseteq Y_{K'}$). So none of the targets in set K' is reached with probability 1 (but it is possible with a positive probability) from a non-terminal spawned off of the path w.

It follows that for a starting non-terminal $T_i \in C'$:

$$\exists\sigma' \in \Psi: \ Pr^{\sigma'}_{T_i}\Big[\Big(\bigcap_{q\in K'} Reach(T_q)\Big) \cap \neg\Box C'\Big] = 1 \ \text{ if and only if}$$

$$\exists\sigma'' \in \Psi: \max_{\langle T_v \in S_{K'} \cup Z_{K'} \cup Y^{t-1}_{K'} \mid \exists T_j \in C', b \in \Gamma^j:\, T_j \xrightarrow{b} T_v\rangle} Pr^{\sigma''}_{T_v}\Big[\bigcap_{q\in K'} Reach(T_q)\Big] = 1$$

The right-hand side of this statement is clearly not satisfied since we already know that $T_v \in S_{K'} \cup Z_{K'} \cup Y^{t-1}_{K'}$ satisfy property $(B)^v_{K'}$.

So it follows that $\forall\sigma' \in \Psi: \ Pr^{\sigma'}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] = Pr^{\sigma'}_{T_i}[(\bigcap_{q\in K'} Reach(T_q)) \cap \neg\Box C'] < 1$.

As for nodes (non-terminals) $T_i \in Y^t_{K'}$ at level t, that do not belong to any SCC, using a similar argument, $\forall\sigma \in \Psi: \ Pr^{\sigma}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] < 1$.

By this inductive argument, it follows that for any non-terminal $T_i \in Y_{K'}$ and for any strategy $\sigma \in \Psi$: $Pr^{\sigma}_{T_i}[\bigcap_{q\in K'} Reach(T_q)] < 1$, i.e., property $(B)^i_{K'}$ is satisfied. □

Now we show that for non-terminals $T_i \in F_{K'}$, when the inner loop for subset $K' \subseteq K$ terminates, the property $(A)^i_{K'}$ is satisfied. We will also construct a witness strategy, under which this property holds for each non-terminal $T_i \in F_{K'}$. Since we have already proved it for non-terminals in set $D_{K'}$, in the following Lemma we refer to the part of set $F_{K'}$ not containing set $D_{K'}$, i.e., to set $F_{K'} = X - S_{K'}$.

Lemma 4.10. *Every non-terminal $T_i \in F_{K'}$ satisfies property $(A)^i_{K'}$.*

Proof. For the rest of this proof denote by $F^0_{K'}$ the initialized set at step II.8. Let us first observe the properties for the non-terminals $T_i \in F_{K'} = X - S_{K'}$. None of the non-terminals is a target non-terminal from set K', i.e., $i \notin K'$. If T_i is of L-form, then:

(L) a non-terminal T_i generates with probability 1 as offspring some non-terminal belonging either to set $F_{K'}$ or to set $D_{K'}$.

If T_i is of M-form, then $\forall a \in \Gamma^i : T_i \xrightarrow{a} T_d,\ T_d \notin D_{K'}$ and:

(M) $\exists a^* \in \Gamma^i : T_i \xrightarrow{a^*} T_j,\ T_j \in F_{K'}$.

If T_i is of Q-form (i.e., $T_i \xrightarrow{1} T_j\, T_r$), then $T_j, T_r \notin D_{K'}$ and:

(Q.0) if $T_i \in F^0_{K'}$, $\exists q \in K'$ such that w.l.o.g. $T_j \in F_{K'} \wedge T_r \in \bar{Z}_{\{q\}}$, (Q.1) otherwise, w.l.o.g. $T_j \in F_{K'}$.

($\mathfrak{P}$) Let us recall that for every $q \in K'$, there is a deterministic static strategy $\sigma'_{\{q\}}$ for the player and a value $b_{\{q\}} > 0$ such that, starting at a non-terminal $T_r \in \bar{Z}_{\{q\}}$, $Pr_{T_r}^{\sigma'_{\{q\}}}[Reach(T_q)] \geq b_{\{q\}}$. Let $b := \min_{q \in K'}\{b_{\{q\}}\} > 0$.

We construct now the non-static witness strategy $\sigma_{K'}$ for the player in the following way. In each generation, there is going to be one non-terminal in the generation that is declared to be a "queen" and the rest of the non-terminals in the generation are called "workers" (we will see the difference between the two labels, especially in the choices of actions). Suppose the initial population is a non-terminal $T_v \in F_{K'}$, declared to be the initial queen.

Consider any finite ancestor history h, that starts at the initial non-terminal $T_v \in F_{K'}$ and all non-terminals in h belong to set $F_{K'}$. Let o denote the non-terminal copy at the end of the ancestor history h. If o is a queen of some L-form non-terminal T_i, then from statement (L) we know that in the next generation the single generated successor child o' is of some non-terminal $T_j \in F_{K'} \cup D_{K'}$. If $T_j \in D_{K'}$, then we use at o' and its subtree of descendants the randomized non-static witness strategy from property $(A)^j_{K'}$ as if the play is starting in o'. If $T_j \in F_{K'}$, then we label o' as the queen in the next generation and use the same strategy $\sigma_{K'}$ (that is currently being described) at it. If o is a queen of some M-form non-terminal T_i, then $\sigma_{K'}$ chooses at o uniformly at random among actions a^* from statement (M) and, hence, in the next generation a single child o' of some non-terminal $T_j \in F_{K'}$ will be generated. Again o' is declared to be the queen in the next generation and the same strategy $\sigma_{K'}$ (currently being described) is used at it. If o is a queen of some Q-form non-terminal T_i (i.e., $T_i \xrightarrow{1} T_j\ T_r$), then there are two cases for the two successor children o' and o'' of non-terminals T_j and T_r, respectively:

- **either** property (Q.0) is satisfied, i.e., $T_i \in F^0_{K'}$, and $T_j \in F_{K'} \wedge T_r \in \bar{Z}_{\{q\}}$, for some target $q \in K'$. Then, in the next generation, we declare o' to be the queen and use the currently described strategy $\sigma_{K'}$ for it. As for the child o'', it is declared to be a worker and the strategy used at the entire subtree, rooted at o'', is some strategy $\sigma'_{\{q'\}}$ (from statement ($\mathfrak{P}$)), where $q' \in K'$ is chosen uniformly at random among all targets $q \in K'$ such that $T_r \in \bar{Z}_{\{q\}}$. The randomization in the strategy of the worker is needed, since non-terminal T_i can belong to more than one set H_q, i.e., T_r can belong to more than one set $\bar{Z}_{\{q\}}$.

- **or** property (Q.1) is satisfied, i.e., $T_i \in F_{K'} - F^0_{K'}$ and w.l.o.g. $T_j \in F_{K'}$. Then, in the next generation, the child o' is again declared to be the queen and the same strategy $\sigma_{K'}$ is used for it, whereas the child o'' is again labelled as a worker, but the strategy for it is irrelevant and so an arbitrary one is chosen for its entire subtree of descendants.

That concludes the description of strategy $\sigma_{K'}$. Now we need to show that, indeed, the randomized non-static strategy $\sigma_{K'}$ is an almost-sure strategy for the player, i.e., that for any $T_i \in F_{K'} : \, Pr^{\sigma_{K'}}_{T_i}[\bigcap_{q \in K'} Reach(T_q)] = 1$.

As previously stated, $F^0_{K'}$ is the initial set $F_{K'}$ at step II.8. Also let $T_{x_1}, T_{x_2}, \ldots, T_{x_t}$ be the non-terminals in set $F_{K'} - F^0_{K'}$ indexed with respect to the time at which they were added to set $F_{K'}$ at step II.9. Let $\gamma := \max_{i \in [n]} |\Gamma^i|$ and let λ be the minimum of $\frac{1}{\gamma}$ and the minimum rule probability in the OBMDP.

Starting at a non-terminal $T_v \in F_{K'}$, consider the sequence of queens. We claim that from any queen with a positive probability $\geq \lambda^n$ in the next $\leq n = |V|$ generations we reach a Q-form queen of a (specific) non-terminal in set $F^0_{K'}$. To show this, we define, for each non-terminal $T_i \in F_{K'}$, a finite "auxiliary" tree $\mathcal{T}_i$, rooted at T_i, which represents why T_i was added to set $F_{K'}$ (i.e., based on steps II.8. and II.9. in the last iteration of the inner loop before step II.11. terminates the inner loop). If $T_i \in F^0_{K'}$, then the tree $\mathcal{T}_i$ is constructed of just a single node (leaf) labelled by T_i. If T_i is of L-form, added at step II.9., then $T_i \rightarrow T_j$, $T_j \in F_{K'}$ (otherwise T_i would have been added to set $D_{K'}$) and the tree $\mathcal{T}_i$ has an edge from its root (labelled by T_i) to a child labelled by T_j (the root of the subtree $\mathcal{T}_j$), for each such $T_j \in F_{K'}$. If T_i is of M-form, added at step II.9., then the tree $\mathcal{T}_i$ has an edge from its root (labelled by T_i) to a child labelled by T_j (the root of the subtree $\mathcal{T}_j$), for every T_j such that $\exists a^* \in \Gamma^i : \, T_i \xrightarrow{a^*} T_j$, $T_j \in F_{K'}$. And if T_i is of Q-form, added at step II.9., then the tree $\mathcal{T}_i$ has an edge from its root (labelled by T_i) to a child labelled by T_j (from property (Q.1)), which is the root of the subtree $\mathcal{T}_j$.

The "auxiliary" tree, just defined, has depth of at most n, since there is a strict order in which the non-terminals entered set $F_{K'}$. Now observe that, if we consider any generation of the play, assuming that the current queen (in this generation) is of some non-terminal $T_i \in F_{K'}$, it can be inductively shown that with a positive probability (at least λ^n) in at most n generations the sequence of queens follows a *specific* root-to-leaf path in $\mathcal{T}_i$. That is because if we are at a queen of a L-form non-terminal T_j (respectively, in node labelled by T_j, which is the root of tree $\mathcal{T}_j$), then in the next generation with probability $\geq \lambda$ the successor queen is of non-terminal $T_{j'} \in F_{K'}$, which is a child of the root of $\mathcal{T}_j$ and is also itself the root of $\mathcal{T}_{j'}$. And if we are at a queen of a M-form non-terminal T_j, then in the next generation (due to the fixed strategy $\sigma_{K'}$) with probability $\geq 1/|\Gamma^j| \geq 1/\gamma \geq \lambda$ the successor queen is of a non-terminal $T_{j_a} \in F_{K'}$, which is a child of the root of $\mathcal{T}_j$ and is also the root of $\mathcal{T}_{j_a}$. And if we are at a queen of a Q-form non-terminal T_j, which is not a leaf in this "auxiliary" tree, then in the next generation with probability 1 the successor queen is of a non-terminal $T_{j'}$, which is the root of $\mathcal{T}_{j'}$ and the unique child of the root of $\mathcal{T}_j$. Since the depth of the "auxiliary" defined tree is at most n, then with probability $\geq \lambda^n$, from a current queen of some non-terminal $T_i \in F_{K'}$, in the next $\leq n$ steps we arrive at a *specific* leaf T_u of the tree $\mathcal{T}_i$, i.e., a queen of non-terminal $T_u \in F^0_{K'}$ is generated.

If somewhere along the sequence of queens, a queen of a L-form non-terminal happens to generate a non-terminal in set $D_{K'}$, then the sequence of queens is actually finite. Therefore, if the sequence of queens is infinite, since it has to follow root-to-leaf paths in the defined "auxiliary" tree, then it follows that with probability 1 infinitely often a queen of a Q-form non-terminal in set $F^0_{K'}$ is observed.

Now consider any $q \in K'$ and any Q-form non-terminal $T_u \in F^0_{K'} \cap H_q$. Since in the subgraph of the dependency graph, induced by $X - S_{K'} = F_{K'}$ (i.e., $G[F_{K'}]$), node T_u is part of a SCC that contains at least one node (non-terminal) from each set $H_{q'}, q' \in K'$, then, along the sequence of queens, from a queen of non-terminal T_u, for any $q' \in K'$ there is a non-terminal $T_{u'} \in F^0_{K'} \cap H_{q'}$ that can be reached as a queen, under the described strategy $\sigma_{K'}$, in at most n generations with a positive probability bounded away from zero (in fact, $\geq \lambda^n$). Note: There is a positive probability, under strategy $\sigma_{K'}$, to exit the particular SCC of T_u. However, under $\sigma_{K'}$ and starting at any non-terminal $T_v \in F_{K'}$, almost-surely the sequence of queens eventually reaches a queen whose non-terminal is in a SCC, C'', in $G[F_{K'}]$ which can only have an outgoing edge to set $D_{K'}$ and where, moreover, for each target in K' there is a branching (Q-form) node in C'' whose "extra" child can hit that target with a positive probability (bounded away from zero).

Hence, starting at a non-terminal $T_v \in F_{K'}$ and under strategy $\sigma_{K'}$, the sequence of queens follows root-to-leaf paths in the defined "auxiliary" tree and, for each $q \in K'$, infinitely often a queen of a Q-form non-terminal from set H_q is observed. And each such queen generates an independent worker, that reaches the respective target non-terminal T_q in a finite expected number of generations with a positive probability bounded away from zero (due to the uniformly at random choice over strategies from statement ($\mathfrak{P}$), for each worker, and due to the fact that the value $b > 0$ from statement ($\mathfrak{P}$) does not depend on the history or the time when the worker is generated). And, more importantly, since the Q-form non-terminals from the sets $F^0_{K'} \cap H_q$ ($q \in K'$) form SCCs in $G[F_{K'}]$, then collectively the independent workers (under their respective strategies) have infinitely often a positive probability bounded away from zero to reach all target non-terminals from set K' in a finite expected number of generations (by Claim 4.11 and by the fact that each independent worker has probability $\geq \frac{b}{k}$ to reach the respective target non-terminal in finite expected number of generations). Hence, all target non-terminals from set K' are reached with probability 1. □

This completes the proof of Theorem 4.7 and the analysis of the almost-sure algorithm. The proof of Lemma 4.10 describes how to construct, for any subset $K' \subseteq K$, the witness strategy $\sigma_{K'}$ for the non-terminals in set $F_{K'}$. These non-static strategies $\sigma_{K'}$ are described as functions that map finite ancestor histories belonging to the controller to distributions over actions available for the current non-terminal of the ancestor history, and can be described in such a form in time $4^k \cdot |\mathcal{A}|^{O(1)}$. □

Recall that we denote by λ the minimum of $\frac{1}{\max_{i \in [n]} |\Gamma^i|}$ and the minimum rule probability in the OBMDP.

Claim 4.11. *If the sequence of queens is not finite, the expected number of generations, starting at a non-terminal $T_v \in F_{K'} - D_{K'}$ and under the strategy $\sigma_{K'}$ constructed in the proof of Lemma 4.10, to observe at least one queen of some non-terminal in each of the sets H_q, $q \in K'$, is $\leq \frac{n}{\lambda^n} \cdot (\ln k + 1)$.*

Proof. Fix strategy $\sigma_{K'}$, constructed in the proof of Lemma 4.10 above. As mentioned in the proof of Lemma 4.10, from a copy of any non-terminal $T_i \in F_{K'} - D_{K'}$, any particular Q-form non-terminal in set $F^0_{K'}$ is reached with probability $\geq \lambda^n$ in the next $\leq$ n generations. Alternatively, for any $T_i \in F_{K'} - D_{K'}$, with probability $\geq \lambda^n$ the sequence of queens in the next $\leq n$ generations of the play follows a specific root-to-leaf path in the associated "auxiliary" tree $\mathcal{T}_i$, defined in the proof of Lemma 4.10.

Let Y_w be a random variable, denoting the number of such root-to-leaf paths (each of length at most n) in the infinite sequence of queens, having already observed at least one queen of a Q-form non-terminal from $w-1$ *different* sets $\langle H_{q_t} \mid q_t \in K', t \in [w-1] \rangle$, to observe a queen of a Q-form non-terminal of a new set $H_{q'}$, $q' \in K'$ (i.e., $q' \neq q_t, \forall t \in [w-1]$). Notice that each Y_w is a geometric random variable and let p_w denote the associated parameter with random variable Y_w. For $l = |K'|$, $Y = \sum_{w=1}^{l} Y_w$ is the total number of root-to-leaf paths in the infinite sequence of queens to observe at least one queen of a non-terminal from each of the sets H_q, $q \in K'$, under strategy $\sigma_{K'}$

and starting at some non-terminal $T_v \in F_{K'} - D_{K'}$. Informally, note that if we think of the sets H_q, $q \in K'$, as l coupons that are to be collected, then this is indeed the famous *coupon collector's problem*.

Denote by $\bar{p}_w$ the probability of observing a queen in the next $\leq n$ generations of a Q-form non-terminal from *any* one of the $w-1$ "collected" sets H_{q_t} ($q_t \in K'$, $t \in [w-1]$). Then clearly:

$$\lambda^n \leq \bar{p}_w \quad \Leftrightarrow \quad p_w \leq 1-\lambda^n, \qquad w \geq 2$$

where to recall $\lambda^n \in (0,1)$ is the least probability of observing a queen in the next $\leq n$ generations of a *particular* Q-form non-terminal from a *particular* set H_q, $q \in K'$, i.e., the least probability of a specific root-to-leaf path in the "auxiliary" tree. Note that the inequality is true only for $w \geq 2$ and that $p_1 = 1$. Then $E[Y_w] = \frac{1}{p_w} \geq \frac{1}{1-\lambda^n}$, for $w \geq 2$, and so:

$$E[Y] = \sum_{w=1}^{l} E[Y_w] \geq 1 + \sum_{w=2}^{l} \frac{1}{1-\lambda^n} = \frac{l-\lambda^n}{1-\lambda^n}$$

For the upper bound on the expectation, notice that each set H_q, $q \in K'$, has cardinality ≥ 1. Then it follows that $p_w \geq (l-w+1)\lambda^n$, for $w \geq 1$, and so that $E[Y_w] = \frac{1}{p_w} \leq \frac{1}{(l-w+1)\lambda^n}$. Then:

$$E[Y] \leq \sum_{w=1}^{l} \frac{1}{(l-w+1)\lambda^n} = \frac{1}{\lambda^n} \sum_{w=1}^{l} \frac{1}{w} = \frac{H_l}{\lambda^n} \leq \frac{H_k}{\lambda^n} < \frac{\ln k + 1}{\lambda^n}$$

The claim follows from the fact that a root-to-leaf path in the "auxiliary" tree is of length at most n.

Note that $\lambda \geq 2^{-poly(|\mathcal{A}|)}$. Then, assuming the sequence of queens is infinite, the expected number of generations, starting at some non-terminal $T_v \in F_{K'} - D_{K'}$ and under the strategy $\sigma_{K'}$ constructed in the proof of Lemma 4.10, to observe at least one queen of a Q-form non-terminal from each of the sets H_q, $q \in K'$, is $\leq 2^{poly(|\mathcal{A}|)} \cdot n \cdot (\ln k + 1)$, i.e., can be exponential in the size of $\mathcal{A}$. □

4.5 Further cases of qualitative multi-objective reachability

In this section we present algorithms for deciding some other cases of qualitative multi-objective reachability problems for OBMDPs, involving certain kinds of Boolean combinations of qualitative reachability and non-reachability queries with respect to given target non-terminals.

4.5.1 $\overset{?}{\exists}\,\sigma \in \Psi :\ \bigwedge_{q\in K} Pr^{\sigma}_{T_i}[Reach(T_q)] < 1$

Proposition 4.12. *There is an algorithm that, given an OBMDP, $\mathcal{A}$, and a set $K \subseteq [n]$ of $k = |K|$ target non-terminals, computes the set $F := \{T_i \in V \mid \exists\sigma \in \Psi : \bigwedge_{q\in K} Pr^{\sigma}_{T_i}[Reach(T_q)] < 1\}$. The algorithm runs in time $k \cdot |\mathcal{A}|^{O(1)}$ and can also compute a randomized static witness strategy σ for the non-terminals in set F.*

Proof. First, as a preprocessing step, for each $q \in K$ we compute the set $W_q := \{T_i \in V \mid \exists\sigma_q \in \Psi : Pr^{\sigma_q}_{T_i}[Reach(T_q)] < 1\}$, together with a single deterministic static strategy σ_q that witnesses the property for every non-terminal in set W_q. This can be done in time $k \cdot |\mathcal{A}|^{O(1)}$, using the algorithms from [ESY18, Proposition 4.1 and Theorem 9.3] for each target T_q, $q \in K$.

Then the Proposition is a direct consequence from the following Claim.

Claim 4.13. $F = \bigcap_{q\in K} W_q$.

Proof. In order to prove the claim, we show the following: $T_i \in \bigcap_{q\in K} W_q$ if and only if $\exists\sigma' \in \Psi :\ \bigwedge_{q\in K} Pr^{\sigma'}_{T_i}[Reach(T_q)] < 1$.

($\Leftarrow$.) Suppose $T_i \notin \bigcap_{q\in K} W_q$, i.e., $T_i \in \bigcup_{q\in K} \overline{W}_q$, where $\overline{W}_q := V - W_q$ for each $q \in K$. Then there exists some $q' \in K$ such that $T_i \in \overline{W}_{q'}$, i.e., $\forall\sigma \in \Psi :\ Pr^{\sigma}_{T_i}[Reach(T_{q'})] = 1$. Clearly, this implies that $\forall\sigma \in \Psi :\ \bigvee_{q\in K} Pr^{\sigma}_{T_i}[Reach(T_q)] = 1$.

($\Rightarrow$.) Suppose that $T_i \in \bigcap_{q\in K} W_q$. Recall that for each $q \in K$ there is a deterministic static witness strategy σ_q for the non-terminals in set W_q. Let σ' be a randomized static strategy for the player defined as follows: in every non-terminal T_j of M-form, let σ' choose uniformly at random among the actions assigned to T_j in each of the deterministic static strategies σ_q, $q \in K$. Hence, for each $q \in K$, there is a positive probability that strategy σ' imitates strategy σ_q. Then, for each target non-terminal T_q ($q \in K$), under σ' and starting at a non-terminal $T_i \in \bigcap_{q\in K} W_q$, it follows that $Pr^{\sigma'}_{T_i}[Reach(T_q)] < 1$. □

The randomized static witness strategy σ for the non-terminals in set F is precisely the strategy σ' constructed in the proof of the Claim above. □

4.5.2 $\overset{?}{\exists}\,\sigma \in \Psi :\ Pr^{\sigma}_{T_i}[\bigcap_{q\in K} Reach(T_q)] < 1$

Proposition 4.14. *There is an algorithm that, given an OBMDP, $\mathcal{A}$, and a set $K \subseteq [n]$ of $k = |K|$ target non-terminals, computes the set $F := \{T_i \in V \mid \exists\sigma \in \Psi :$*

$Pr^{\sigma}_{T_i}[\bigcap_{q \in K} Reach(T_q)] < 1\}$. The algorithm runs in time $k \cdot |\mathcal{A}|^{O(1)}$ and can also compute a deterministic static witness strategy σ for a given starting non-terminal $T_i \in F$.

Proof. First, as a preprocessing step, for each $q \in K$ we compute the set $W_q := \{T_i \in V \mid \exists \sigma_q \in \Psi : Pr^{\sigma_q}_{T_i}[Reach(T_q)] < 1\}$, together with a single deterministic static strategy σ_q that witnesses the property for every non-terminal in set W_q. This can be done in time $k \cdot |\mathcal{A}|^{O(1)}$, using the algorithms from [ESY18, Proposition 4.1 and Theorem 9.3] for each target T_q, $q \in K$.

Then the Proposition is a direct consequence from the claim that $F = \bigcup_{q \in K} W_q$. To see this claim, note that $T_i \in \bigcup_{q \in K} W_q$ if and only if there exists $\sigma' \in \Psi$ and some $q \in K$ such that $Pr^{\sigma'}_{T_i}[Reach(T_q)] < 1$ (by definition of the W_q, $q \in K$ sets). Then the claim follows directly from Proposition 2.2(2.).

For each $T_i \in F$, select some $q \in K$, such that $T_i \in W_q$, and, starting at a non-terminal T_i, let the witness strategy σ act exactly as the deterministic static strategy σ_q. □

Consider the following two examples of OBMDPs consisting of non-terminals $\{M, T, T', L, R_1, R_2\}$ and target non-terminals R_1 and R_2. M is the only controlled non-terminal. The examples provide a good idea of the difference between the objectives in Propositions 4.12 and 4.14.

Example 4.2.

$$M \xrightarrow{a} T \qquad T \xrightarrow{1} L\,R_1 \qquad L \xrightarrow{1/2} \varnothing$$

$$M \xrightarrow{b} T' \qquad T' \xrightarrow{1} R_1\,R_2 \qquad L \xrightarrow{1/2} R_2$$

There exists a deterministic static witness strategy σ' such that $Pr^{\sigma'}_{M}[Reach(R_1) \cap Reach(R_2)] < 1$, namely, starting at a non-terminal M, let the player choose deterministically action a. Thus, the probability of observing both target non-terminals in the generated play (tree) is $1/2$. However, notice that for any strategy σ, starting at non-terminal M, target non-terminal R_1 is reached with probability 1. That is, $\forall \sigma \in \Psi : \bigvee_{q \in \{1,2\}} Pr^{\sigma}_{M}[Reach(R_q)] = 1$.

Example 4.3.

$$M \xrightarrow{a} T \qquad T \xrightarrow{1} L\,R_1 \qquad L \xrightarrow{1/2} R_1$$

$$M \xrightarrow{b} T' \qquad T' \xrightarrow{1} L\,R_2 \qquad L \xrightarrow{1/2} R_2$$

There exists a static strategy σ' such that $\bigwedge_{q\in\{1,2\}} Pr^{\sigma'}_{M}[Reach(R_q)] < 1$, but the strategy needs to randomize, otherwise a deterministic choice in non-terminal M will generate a target non-terminal immediately in the next generation. Note that the same strategy σ' (although a deterministic one suffices) also guarantees $Pr^{\sigma'}_{M}[Reach(R_1) \cap Reach(R_2)] < 1$.

4.5.3 $\overset{?}{\exists} \sigma \in \Psi : \bigwedge_{q\in K} Pr^{\sigma}_{T_i}[Reach(T_q)] > 0$

Proposition 4.15. *There is an algorithm that, given an OBMDP, $\mathcal{A}$, and a set $K \subseteq [n]$ of $k = |K|$ target non-terminals, computes the set $F := \{T_i \in V \mid \exists \sigma \in \Psi : \bigwedge_{q\in K} Pr^{\sigma}_{T_i}[Reach(T_q)] > 0\}$. The algorithm runs in time $O(k \cdot |V|^2)$ and can also compute a randomized static witness strategy σ for the non-terminals in set F.*

Proof. First, for each $q \in K$, we compute the attractor set of target non-terminal T_q with respect to the dependency graph $G = (U, E)$, $U = V$, of $\mathcal{A}$. That is, for each $q \in K$, we compute the set $Attr(T_q)$ as the limit of the following sequence $\langle Attr_t(T_q) \mid t \geq 0\rangle$:

$$Attr_0(T_q) = \{T_q\}$$
$$Attr_t(T_q) = Attr_{t-1}(T_q) \cup \{T_i \in V \mid \exists\, T_j \in Attr_{t-1}(T_q) \text{ s.t. } (T_i, T_j) \in E\}$$

In other words, $Attr(T_q)$ is the set of nodes in G (or equivalently, non-terminals in $\mathcal{A}$) that have a directed path to the target node (non-terminal) T_q in the dependency graph G. For each $q \in K$, such a set can be computed in time $O(|V|^2)$. So all k attractor sets (one for each target non-terminal $T_q, q \in K$) can be computed in time $O(k \cdot |V|^2)$. The Proposition is a direct consequence from the following Claim.

Claim 4.16. *$F = \bigcap_{q\in K} Attr(T_q)$.*

Proof. To prove the Claim, we need to show that $T_i \in \bigcap_{q\in K} Attr(T_q)$ if and only if $\exists \sigma' \in \Psi : \bigwedge_{q\in K} Pr^{\sigma'}_{T_i}[Reach(T_q)] > 0$.
($\Leftarrow$.) Suppose that $T_i \notin \bigcap_{q\in K} Attr(T_q)$, i.e., there exists some $q' \in K$ such that $T_i \notin Attr(T_{q'})$. This implies that in the dependency graph G there is even no path from T_i to $T_{q'}$. Therefore, regardless of strategy σ for the player, $Pr^{\sigma}_{T_i}[Reach(T_{q'})] = 0$ and hence, $\forall \sigma \in \Psi : \bigvee_{q\in K} Pr^{\sigma}_{T_i}[Reach(T_q)] = 0$.

($\Rightarrow$.) Suppose that $T_i \in \bigcap_{q\in K} Attr(T_q)$. Let σ' be the randomized static strategy such that in every non-terminal $T_j \in V$ of M-form it chooses uniformly at random an action among its set of actions Γ^j. For each $q \in K$, in the dependency graph G there is a directed path from T_i to T_q. Then under the described strategy σ', starting at a non-terminal T_i, there is a positive probability to generate any of the target non-terminals $\{T_q \mid q \in K\}$, because there is a positive probability for a path in the play (tree) to follow the directed path in G from T_i to T_q, for any $q \in K$.

Denote by λ the minimum of $\frac{1}{\max_{j\in[n]}|\Gamma^j|}$ and the minimum probability among the probabilistic rules of $\mathcal{A}$. Then, in fact, for each $q \in K$, under σ' there is a probability $\geq \lambda^n$ to generate a copy of target non-terminal T_q in the next $\leq n$ generations, i.e., $\bigwedge_{q\in K} Pr_{T_i}^{\sigma'}[Reach(T_q)] \geq \lambda^n > 0$. □

The randomized static witness strategy σ for the non-terminals in set F is the strategy σ' constructed in the proof of the Claim above. □

4.5.4 $\overset{?}{\exists}\, \sigma \in \Psi : \; Pr_{T_i}^{\sigma}[\bigcap_{q\in K} Reach^{\complement}(T_q)] \triangle \{0,1\}$

Now let us consider the qualitative cases of multi-objective reachability where for a given OBMDP and a given set $K \subseteq [n]$ of target non-terminals, the aim is to compute those non-terminals $T_i \in V$ such that $\exists \sigma \in \Psi : \; Pr_{T_i}^{\sigma}[\bigcap_{q\in K} Reach^{\complement}(T_q)] \triangle \{0,1\}$, where $\triangle := \{<,=,>\}$.

First, due to the fact that the complement of the set (of plays) $\bigcap_{q\in K} Reach^{\complement}(T_q)$ is the set (of plays) $\bigcup_{q\in K} Reach(T_q)$, we give the following Lemma to show that this complement objective reduces to the objective of reachability of a single target non-terminal in a slightly modified OBMDP.

Lemma 4.17. *There is an algorithm that, given an OBMDP, $\mathcal{A}$, and a set $K \subseteq [n]$ of $k = |K|$ target non-terminals $\{T_q \in V^{\mathcal{A}} \mid q \in K\}$, runs in linear time $O(|\mathcal{A}|)$ and outputs another OBMDP, $\mathcal{A}'$, with a single target non-terminal T_f, such that for any $T_i \in V^{\mathcal{A}} - \{T_q \in V^{\mathcal{A}} \mid q \in K\} = V^{\mathcal{A}'} - \{T_f\}$ and any strategy $\sigma \in \Psi^{\mathcal{A}}$, there exists a strategy $\sigma' \in \Psi^{\mathcal{A}'}$ such that $Pr_{T_i}^{\sigma,\mathcal{A}}[\bigcup_{q\in K} Reach(T_q)] = Pr_{T_i}^{\sigma',\mathcal{A}'}[Reach(T_f)]$.*

Proof. Consider the OBMDP, $\mathcal{A}'$, obtained from OBMDP, $\mathcal{A}$, by adding a new purely probabilistic target non-terminal T_f with a single rule $T_f \xrightarrow{1} \varnothing$, removing all target non-terminals $\{T_q \in V^{\mathcal{A}} \mid q \in K\}$ and their associated rules, and replacing any occurrence of a non-terminal $T_q \in V^{\mathcal{A}}$, $q \in K$, on the right-hand side of some rule with non-terminal T_f. Hence, $V^{\mathcal{A}'} = (V^{\mathcal{A}} \cup \{T_f\}) - \{T_q \in V^{\mathcal{A}} \mid q \in K\}$. Clearly, for any $T_q \in V^{\mathcal{A}}$, with $q \in K$ and for any $\sigma \in \Psi^{\mathcal{A}}$, $Pr_{T_q}^{\sigma,\mathcal{A}}[\bigcup_{q'\in K} Reach(T_{q'})] = 1$. Also, for $T_f \in V^{\mathcal{A}'}$ and for any $\sigma' \in \Psi^{\mathcal{A}'}$, $Pr_{T_f}^{\sigma',\mathcal{A}'}[Reach(T_f)] = 1$.

Observe that for any play (tree) $\mathcal{T}$ in $\mathcal{A}$, there is a play $\mathcal{T}'$ in $\mathcal{A}'$ such that any copy o of a non-terminal $T_q \in V^{\mathcal{A}}$, $q \in K$, in $\mathcal{T}$ is replaced in $\mathcal{T}'$ by a copy of non-terminal T_f and the subtree of descendants of o is non-existent in $\mathcal{T}'$.

Now consider any starting non-terminal $T_u \in V^{\mathcal{A}} - \{T_q \in V^{\mathcal{A}} \mid q \in K\} = V^{\mathcal{A}'} - \{T_f\}$.

Let $\sigma \in \Psi^{\mathcal{A}}$ be any strategy for the player in $\mathcal{A}$. Define strategy $\sigma' \in \Psi^{\mathcal{A}'}$ in $\mathcal{A}'$ in the following way: for each non-terminal $T_i \in V^{\mathcal{A}'} - \{T_f\}$, strategy σ' behaves exactly like σ for all ancestor histories ending in T_i, and for non-terminal T_f strategy σ' acts arbitrarily in all ancestor histories ending in T_f since it is irrelevant. Note that, due to the construction of $\mathcal{A}'$ and σ', if a play (tree) $\mathcal{T}$, generated under strategy σ, belongs to the set (objective) $\bigcup_{q \in K} Reach(T_q)$ in $\mathcal{A}$, then in $\mathcal{A}'$ under σ' the corresponding unique play $\mathcal{T}'$ (as described above) belongs to the set (objective) $Reach(T_f)$. Furthermore, all plays $\mathcal{T}$ in $\mathcal{A}$ with the same corresponding play $\mathcal{T}'$ in $\mathcal{A}'$ have a combined probability, of being generated under σ, equal to the probability of $\mathcal{T}'$ being generated under σ' in $\mathcal{A}'$. Hence, $Pr_{T_u}^{\sigma,\mathcal{A}}[\bigcup_{q \in K} Reach(T_q)] = Pr_{T_u}^{\sigma',\mathcal{A}'}[Reach(T_f)]$. But σ was an arbitrary strategy.

For the opposite direction, let $\sigma' \in \Psi^{\mathcal{A}'}$ be any strategy for the player in $\mathcal{A}'$. Define $\sigma \in \Psi^{\mathcal{A}}$ to be the strategy in $\mathcal{A}$ such that, for each non-terminal $T_i \in V^{\mathcal{A}} - \{T_q \in V^{\mathcal{A}} \mid q \in K\}$, acts the same as σ' in all ancestor histories ending in T_i; and for each non-terminal $T_q \in V^{\mathcal{A}}$, $q \in K$, the strategy σ acts arbitrarily in all ancestor histories ending in T_q as it is irrelevant. Then, for any play $\mathcal{T}' \in Reach(T_f)$ in $\mathcal{A}'$ under strategy σ', there is at least one play $\mathcal{T} \in \bigcup_{q \in K} Reach(T_q)$ in $\mathcal{A}$ under strategy σ, such that for any copy of non-terminal T_f in tree $\mathcal{T}'$ there is a copy of some non-terminal $T_q \in V^{\mathcal{A}}$, $q \in K$, at the corresponding position in tree $\mathcal{T}$. But note that the probability of generating $\mathcal{T}'$ in $\mathcal{A}'$ under σ' is equal to the sum of probabilities of generating all such corresponding plays $\mathcal{T}$ in $\mathcal{A}$ under σ. Hence, $Pr_{T_u}^{\sigma',\mathcal{A}'}[Reach(T_f)] = Pr_{T_u}^{\sigma,\mathcal{A}}[\bigcup_{q \in K} Reach(T_q)]$. But σ' was an arbitrary strategy. □

We now present a Proposition that deals with all four qualitative questions for the (set of plays) objective $\bigcap_{q \in K} Reach^{\complement}(T_q)$ for a given set $K \subseteq [n]$ of target non-terminals.

Proposition 4.18. *There is a P-time algorithm that, given an OBMDP, $\mathcal{A}$, and a set $K \subseteq [n]$ of $k = |K|$ target non-terminals, computes the set $F := \{T_i \in V \mid \exists \sigma \in \Psi : Pr_{T_i}^{\sigma}[\bigcap_{q \in K} Reach^{\complement}(T_q)] \triangle \{0,1\}\}$, where $\triangle := \{<,=,>\}$. The algorithm can also compute a deterministic witness strategy σ for the non-terminals in set F.*

Proof. We can rephrase the question of whether $\exists \sigma \in \Psi^{\mathcal{A}} : Pr_{T_i}^{\sigma,\mathcal{A}}[\bigcap_{q \in K} Reach^{\complement}(T_q)] \triangle x$, where $x \in \{0,1\}$, accordingly into the form $\exists \sigma \in \Psi^{\mathcal{A}} : Pr_{T_i}^{\sigma,\mathcal{A}}[\bigcup_{q \in K} Reach(T_q)] \triangle_{\complement} 1 - x$, where $\triangle_{\complement}$ is $<,=,>$ if $\triangle$ is $>,=,<$, respectively. And as a consequence of Lemma 4.17, there exists a modified OBMDP, $\mathcal{A}'$, with a single target non-terminal T_f such that $\mathcal{A}'$ is computable in linear time and the following is true: $\exists\sigma \in \Psi^{\mathcal{A}} : Pr_{T_i}^{\sigma,\mathcal{A}}[\bigcup_{q\in K} Reach(T_q)] \triangle_{\complement} 1-x$ if and only if $\exists\sigma' \in \Psi^{\mathcal{A}'} : Pr_{T_i}^{\sigma',\mathcal{A}'}[Reach(T_f)] \triangle_{\complement} 1-x$.

For the case of $1-x=0$, by [ESY18, Proposition 4.1], there is a P-time algorithm to compute the set $F^{\mathcal{A}'}$ of non-terminals T_i in $\mathcal{A}'$ and a deterministic static witness strategy $\sigma' \in \Psi^{\mathcal{A}'}$ such that $T_i \in F^{\mathcal{A}'}$ are precisely the non-terminals that satisfy the property $Pr_{T_i}^{\sigma',\mathcal{A}'}[Reach(T_f)] \triangle_{\complement} 0$.

For the case of $1-x=1$ and $\triangle_{\complement}$ equal to $<$ (respectively, $=$), by [ESY18, Proposition 4.1 and Theorems 9.3, 9.4], there is again a P-time algorithm to compute the set $F^{\mathcal{A}'}$ of non-terminals T_i in $\mathcal{A}'$ and a deterministic static (respectively, non-static) witness strategy $\sigma' \in \Psi^{\mathcal{A}'}$ such that $T_i \in F^{\mathcal{A}'}$ are precisely the non-terminals that satisfy the property $Pr_{T_i}^{\sigma',\mathcal{A}'}[Reach(T_f)] < 1$ (respectively, $Pr_{T_i}^{\sigma',\mathcal{A}'}[Reach(T_f)] = 1$).

Now for the qualitative decision questions where tuple $(\triangle_{\complement}, 1-x)$ is equal to $(=,0)$ or $(<,1)$, let $F = F^{\mathcal{A}} := F^{\mathcal{A}'}$; and where tuple $(\triangle_{\complement}, 1-x)$ is equal to $(>,0)$ or $(=,1)$, let $F = F^{\mathcal{A}} := (F^{\mathcal{A}'} - \{T_f\}) \cup \{T_q \in V^{\mathcal{A}} \mid q \in K\}$. By the proof of Lemma 4.17, from the deterministic (non-)static witness strategy $\sigma' \in \Psi^{\mathcal{A}'}$ in $\mathcal{A}'$ for the starting non-terminals from set $F^{\mathcal{A}'}$ we can obtain a corresponding deterministic (non-)static witness strategy $\sigma \in \Psi^{\mathcal{A}}$ in $\mathcal{A}$ for the starting non-terminals from set $F - \{T_q \in V^{\mathcal{A}} \mid q \in K\}$. As for each non-terminal $T_q \in V^{\mathcal{A}}, q \in K$, let strategy σ make deterministically and statically an arbitrary choice of action from the action set Γ^q (in the case if T_q is of M-form), since if $T_q \notin F$ then strategy is irrelevant at T_q and if $T_q \in F$ then the property holds for any choice of the strategy in T_q. □

4.5.5 $\overset{?}{\exists}\sigma \in \Psi : \bigwedge_{q\in K} Pr_{T_i}^{\sigma}[Reach(T_q)] = 0$

Proposition 4.19. *There is a P-time algorithm that, given an OBMDP, $\mathcal{A}$, and a set $K \subseteq [n]$ of $k = |K|$ target non-terminals, computes the set $F := \{T_i \in V \mid \exists\sigma \in \Psi : \bigwedge_{q\in K} Pr_{T_i}^{\sigma}[Reach(T_q)] = 0\}$. The algorithm can also compute a deterministic static witness strategy σ for the non-terminals in set F.*

Proof. Note that the question of deciding whether there exists a strategy $\sigma \in \Psi$ for the player such that $\bigwedge_{q\in K} Pr_{T_i}^{\sigma}[Reach(T_q)] = 0$ can be rephrased as asking whether there exists a strategy $\sigma \in \Psi$ such that $\bigwedge_{q\in K} Pr_{T_i}^{\sigma}[Reach^{\complement}(T_q)] = 1$. By Proposition 2.2(1.), we already know that it is equivalent to ask instead whether there exists a strategy $\sigma \in \Psi$ such that $Pr_{T_i}^{\sigma}[\bigcap_{q\in K} Reach^{\complement}(T_q)] = 1$. Hence, $F = \{T_i \in V \mid \exists\sigma \in \Psi :$

$Pr^{\sigma}_{T_i}[\bigcap_{q \in K} Reach^{\complement}(T_q)] = 1\}$. And by Proposition 4.18, there is a P-time algorithm to compute the set F and to compute a deterministic static witness strategy σ for the non-terminals in set F. □

We leave open the decidability of general Boolean combinations of arbitrary qualitative reachability and non-reachability queries.

Chapter 5

Conclusions and Future Work

In this thesis we have studied two models of infinite-state stochastic processes, namely Branching Processes and Ordered Branching Processes, where one can naturally view the latter as a crossover model between Branching Processes and Stochastic Context-Free Grammars. In particular, OBPs borrow the simultaneous expansion of rules, generation by generation, from BPs, while also borrowing from SCFGs the fact that there is an ordering among the children generated by any rule and, hence, an ordering of the non-terminals in the generated tree.

To sum up the main results, this thesis included the first study of the (single-target) reachability objective for the concurrent game generalization of BPs. We showed that BCSGs are determined, i.e. have a value, and showed that computing the non-reachability values for a BCSG is equivalent to computing the Greatest Fixed Point of a corresponding system of equations (called a minimax-PPS), by extending known results for the subclass of BSSGs. We also showed that the qualitative almost-sure and limit-sure reachability problems do not coincide in the case of branching concurrent games, and gave the first polynomial time algorithms for both almost-sure and limit-sure reachability in BCSGs. The proofs of the algorithms showed how to compute an almost-sure strategy (respectively, a limit-sure strategy for a given desired error $\varepsilon > 0$) for the player maximizing the reachability probability, or alternatively, a spoiling strategy for the player minimizing the reachability probability if almost-sure (respectively, limit-sure) reachability is not satisfied. Moreover, in the interest of the study of the reachability objective for branching processes being complete in this thesis, we showed that analogous complexity results from past papers on a related line of work (on Recursive models) apply for reachability in BCSGs, thus showing PSPACE to be an upper bound for both quantitative reachability decision and approximation questions for BCSGs and PosSLP to be a lower bound for the quantitative reachability decision questions for BCSGs (even for the purely probabilistic BPs). These are the best bounds we know so far. We also showed that computing the optimal reachability probabilities in a minimizing BMDP (equivalently, computing the GFP of a maxPPS) is in the complexity class, FIXP, which captures search problems that can be rephrased as fixed-point problems.

Furthermore, this thesis included the first look on multi-objective reachability on branching processes, and to be more precise, on the OBPs model. We showed that qualitative multi-objective reachability (particularly, the qualitative problem of multi-target reachability) in OBMDPs is in general NP-hard, when the number of given target non-terminals is unbounded. We also demonstrated that for OBMDPs, unlike in the case of single-target reachability, the almost-sure and limit-sure multi-target reachability problems do not coincide and provided algorithms for both problems that are fixed-parameter tractable with respect to the number k of target non-terminals (i.e., that run in time polynomial in the size of the OBMDP and exponential in k). We also studied for OBMDPs other certain Boolean combinations of qualitative reachability and non-reachability queries with respect to different given target non-terminals, providing efficient algorithmic results for their decidability. In the proofs of all the given algorithms, we showed how to construct the corresponding desired witness strategy for the player in the OBMDP.

5.1 Open problems & Future work

The following is a list of open problems that are suggested as a follow-up study to the work presented in this thesis.

1. It remains open the question of how much the PSPACE upper bounds for the approximation of (non)-reachability values in BCSGs, equivalently approximation of the GFP of an associated minimax-PPS, can be improved. We do not yet know any lower bounds for this problem.

2. Furthermore, recall that in Section 3.6 we showed that computing the optimal reachability probabilities in minimizing BMDPs, equivalently, computing the GFP of a maxPPS, is in FIXP. We leave open the questions of whether computing the optimal reachability probabilities in maximizing BMDPs (equivalently, computing the GFP of a minPPS) and computing the reachability values in BSSGs (equivalently, computing

the GFP of a min-max-PPS) is also in FIXP, and whether approximating the reachability values in BCSGs (equivalently, approximating the GFP of a minimax-PPS) is in FIXP_a.

Note that it has been shown in an unpublished manuscript ([ESY14]) that approximating the extinction values in BCSGs (equivalently, approximating the LFP of a minimax-PPS) is in FIXP_a. It is plausible that similar techniques may prove the inclusion in FIXP_a of the problem of approximating the GFP of a minimax-PPS, but currently this remains an open problem.

3. The decidability of arbitrary Boolean combinations of qualitative reachability and non-reachability queries over different given target non-terminals in OBMDPs remains open. We studied certain cases of the qualitative multi-objective reachability, with the almost-sure and limit-sure multi-target reachability problems being the most interesting to study and the more important on the journey to a complete analysis of arbitrary qualitative questions.

Also, it would be interesting to extend the qualitative multi-objective reachability results, that we provided for OBMDPs, to Ordered Branching Simple (turn-based) Stochastic Games.

4. Furthermore, we leave open (both the decision and approximation) *quantitative* multi-objective reachability questions for OBMDPs. The goal of the quantitative problem is to optimize each of the respective probabilities that the generated tree satisfies each of several given reachability objectives. Clearly, there may be trade-offs between the different objectives. That is, increasing the probability of one of the objectives may result in decreasing the probability of another objective, or in other words, satisfying one objective with a high probability may result in satisfying another with a low probability. That is why in the presence of k objectives, one can be interested in finding vectors of probabilities, $p = (p_1, p_2, \dots, p_k)$, such that there is a strategy for the controller where, for each $i \in [k]$, the i-th objective is guaranteed to be achieved with probability $\geq p_i$. In other words, one may be interested in computing (or approximating) the *trade-off curve* (also called the *Pareto curve*) of optimal probabilities with which the different reachability objectives can be achieved. To be more precise, the Pareto curve is the set of all achievable vectors p of probabilities such that there is no vector $p' \neq p$ where $p' \geq p$ (coordinate-wise). And since it may be computationally expensive to construct the exact curve, often the focus is on approximating the curve.

An ε-approximation for the Pareto curve is the set P_ε of all achievable vectors such that for each achievable vector t there is a vector $p \in P_\varepsilon$ where $(1+\varepsilon)p \geq t$.

5. For (O)BMDPs with a single target, we have Bellman optimality equations whose (greatest) solution captures the optimal (single-target) non-reachability probabilities ([ESY18]), but we do not yet have Bellman optimality equations for multi-objective (non-)reachability. What is more, multi-objective reachability for finite-state MDPs can be characterized as multi-objective linear programming ([EKVY08]). But we do not yet know how to characterize multi-objective reachability for (O)BMDPs as multi-objective mathematical programming, and we believe this is a very promising approach to explore. It may also imply complexity bounds for the quantitative problems.

6. We have shown that under the objective of single-target reachability the OBMDP and BMDP models are equivalent. However, the equivalence is not yet evident under multi-objective reachability. The reason for this is the following. In the BMDP model there is no ordering among the children generated by a rule (the set of offsprings in a rule is a multi-set over the types). Nevertheless, the histories that the controller's strategy maps to distributions on actions are entire finite trees, not just information about the ancestors of a node (i.e., not just what we called an ancestor history). That is, in BMDPs the strategy has at its disposal the entire finite tree up to the "current generation" in the process, together with all the actions chosen and probabilistic rules applied in all previous generations.

Recall that in Section 2.4 we also defined a general strategy for OBMDPs to have as a history the entire finite tree up to the current generation. There is no "good" or "suitable" definition of a strategy for the models of branching processes. There are many variations of the type of history that is provided to the strategy and each one can bring different advantages and disadvantages to the objectives we study in this thesis. We have utilized two natural ways to define the notion of a history for the strategy, but for others there may be another more natural definition. What is interesting in the two variations of a history that we have provided, is that we already showed that the more general notion of a history (i.e., strategy having the information of the entire finite tree up to the current generation) is not more powerful than an ancestor history for OBMDPs for the (single-target) reachability objective, due to the advantage of having ordering in the tree. But there is a difference, not investigated yet, for multi-objective reachability.

Recall that Example 2.1 showed that the almost-sure and limit-sure multi-target reachability problems in OBMDPs do not coincide. However, if the example is an OBMDP where the strategy is allowed to have the entire finite tree up to the current generation as a history, then in that particular example both almost-sure and limit-sure multi-target reachability is satisfied. That is, there is a strategy that, starting in a non-terminal M, guarantees to reach both targets R_1 and R_2 with probability 1. To see this, consider the following deterministic strategy σ': in every generation $t \geq 0$, if R_1 has not yet occurred anywhere in the history tree, then for the unique non-terminal M in the current generation choose (deterministically) action a, yielding a child M and another child A. (Note that each non-terminal A has a $1/2$ chance of having a child R_1.) If, on the other hand, the history tree already contains R_1, then (deterministically) choose action b at the (at most one) non-terminal M in the current generation. It is easy to check that this strategy guarantees that both targets R_1 and R_2 will occur in the play, with probability 1, in a finite expected number of generations.

However, observe that there is *no static* (not even randomized) almost-sure strategy for reaching both target non-terminals. If a static strategy puts any positive probability on action b, then the probability of reaching R_1 is strictly less than 1. Otherwise, the probability of reaching R_2 is 0. (On the other hand, note that there is a family of limit-sure *static* strategies: for each $\varepsilon > 0$, let σ'_ε put probability $1-\varepsilon$ on action a and probability ε on action b. In the limit, as $\varepsilon \to 0$, the probability of generating both R_1 and R_2 in the play approaches 1.)

We believe that almost-sure and limit-sure multi-target reachability problems do coincide for (O)BMDPs with the more general notion of a history for the strategy (where the history is the entire finite tree up to the current generation and not just information about the ancestors of a node), and, in fact, we believe that the algorithm we gave for the limit-sure case in Section 4.3 is sufficient to be the algorithm for qualitative multi-target reachability in such (O)BMDPs. However, we should point out that this is a promising hypothesis to investigate in the future and the problem of qualitative multi-target reachability in (O)BMDPs with the more general notion of a history for the strategy remains an open problem.

As you can see from the open problems posed above, the study contained in this thesis spawns certainly a vast line of further research that can be investigated.

Bibliography

[AB09] S. Arora and B. Barak. *Computational Complexity: A Modern Approach.* Cambridge University Press, 2009.

[ABKPM09] E. Allender, P. Bürgisser, J. Kjeldgaard-Pedersen, and P. B. Miltersen. On the Complexity of Numerical Analysis. *SIAM J. on Computing*, 38(5):1987–2006, 2009. (Conference version in CCC'06.).

[BBFK08] T. Brázdil, V. Brožek, V. Forejt, and A. Kučera. Reachability in Recursive Markov Decision Processes. *Inf. & Comp.*, 206(5):520–537, 2008.

[BBKO11] T. Brázdil, V. Brožek, A. Kučera, and J. Obdržálek. Qualitative Reachability in Stochastic BPA games. *Inf. & Comp.*, 209(8):1160–1183, 2011.

[Bea13] Bozic and et. al. Evolutionary dynamics of cancer in response to targeted combination therapy. *eLife*, 2:e00747, 2013.

[BEKK13] T. Brázdil, J. Esparza, S. Kiefer, and A. Kučera. Analyzing probabilistic pushdown automata. *Formal Methods Syst. Des.*, 43(2):124–163, 2013.

[BKL14] R. Bonnet, S. Kiefer, and A. W. Lin. Analysis of Probabilistic Basic Parallel Processes. In *Proc. of 17th FoSSaCS*, pages 43–57, 2014.

[BKN$^+$19] N. Balaji, S. Kiefer, P. Novotný, G. A. Pérez, and M. Shirmohammadi. On the Complexity of Value Iteration. In *Proc. of 46th ICALP*, pages 102:1–102:15, 2019.

[Blo91] J. Blomer. Computing Sums of Radicals in Polynomial Time. In *Proc. of 32nd IEEE Symposium on Foundations of Computer Science*, pages 670–677, 1991.

[BSS89] L. Blum, M. Shub, and S. Smale. On a Theory of Computation and Complexity over the Real Numbers: NP-completeness, Recursive Functions and Universal Machines. *American Mathematical Society*, 21(1):1–46, 1989.

[Can88] J. Canny. Some Algebraic anf Geometric Computations in PSPACE. In *Proc. of 20th STOC*, 1988.

[CDK12] T. Chen, K. Dräger, and S. Kiefer. Model Checking Stochastic Branching Processes. In *Proc. of 37th MFCS*, pages 271–282, 2012.

[CH14] K. Chatterjee and M. Henzinger. Efficient and Dynamic Algorithms for Alternating Büchi Games and Maximal End-Component Decomposition. *J. of the ACM*, 61(3):15:1–15:40, 2014.

[CHIJ17] K. Chatterjee, K. A. Hansen, and R. Ibsen-Jensen. Strategy Complexity of Concurrent Safety Games. In *Proc. of 42nd MFCS*, 2017.

[Chu01] K. L. Chung. *A Course in Probability Theory*. Academic Press, 2001.

[Con92] A. Condon. The Complexity of Stochastic Games. *Inf. & Comp.*, 96(2):203–224, 1992.

[Con93] A. Condon. On Algorithms for Simple Stochastic Games. *Advances in Computational Complexity Theory, DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 13:51–73, 1993.

[CY98] C. Courcoubetis and M. Yannakakis. Markov decision processes and regular events. *IEEE Transactions on Automatic Control*, 43(10):1399–1418, 1998.

[dAHK07] L. de Alfaro, T. A. Henzinger, and O. Kupferman. Concurrent Reachability Games. *Theoretical Computer Science*, 386(3):188–217, 2007. (Conference version in FOCS'98.).

[DEKM98] R. Durbin, S. R. Eddy, A. Krogh, and G. Mitchison. *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*. Cambridge University Press, 1998.

[DR05] E. Denardo and U. Rothblum. Totally expanding multiplicative systems. *Linear Algebra and its Applications*, 406:142–158, 2005.

[EGK10] J. Esparza, A. Gaiser, and S. Kiefer. Computing least fixed points of probabilistic systems of polynomials. In *Proc. of 27th STACS*, pages 359–370, 2010.

[EKM06] J. Esparza, A. Kucera, and R. Mayr. Model Checking Probabilistic Pushdown Automata. *Logical Methods in Computer Science*, 2(1):1–31, 2006.

[EKVY08] K. Etessami, M. Kwiatkowska, M. Y. Vardi, and M. Yannakakis. Multi-Objective Model Checking of Markov Decision Processes. *Logical Methods on Computer Science*, 4(4), 2008. (Conference version in TACAS'07.).

[EM20] K. Etessami and E. Martinov. Qualitative Multi-Objective Reachability for Ordered Branching MDPs. In *Proc. of 14th Int. Conf. on Reachability Problems (RP)*, 2020. (All references are to the full preprint arXiv:2008.10591).

[EMSY19] K. Etessami, E. Martinov, A. Stewart, and M. Yannakakis. Reachability for Branching Concurrent Stochastic Games. In *Proc. of 46th ICALP*, 2019. (All references are to the full preprint arXiv:1806.03907).

[EPRY20] K. Etessami, C. H. Papadimitriou, A. Rubinstein, and M. Yannakakis. Tarski's Theorem, Supermodular Games, and the Complexity of Equilibria. In *Proc. of 11th ITCS*, pages 18:1–18:19, 2020.

[ESY14] K. Etessami, A. Stewart, and M. Yannakakis. NOTE: Approximating the termination probabilities for RMCs, and the extinction values for Branching Stochastic Games, is in FIXP. unpublished manuscript, 2014.

[ESY17] K. Etessami, A. Stewart, and M. Yannakakis. A Polynomial-time Algorithm for Computing Extinction Probabilities of Multi-type Branching Processes. *SIAM J. on Computing*, 46(5):1515–1553, 2017. (Incorporates part of the work of a conference paper in STOC'12.).

[ESY18] K. Etessami, A. Stewart, and M. Yannakakis. Greatest Fixed Points of Probabilistic Min/Max Polynomial Equations, and Reachability for Branching Markov Decision Processes. *Inf. & Comp.*, 261(2):355–382, 2018. (Conference version in ICALP'15.).

[ESY20] K. Etessami, A. Stewart, and M. Yannakakis. Polynomial Time Algorithms for Branching Markov Decision Processes and Probabilistic Min(Max) Polynomial Bellman Equations. *Math. Oper. Res.*, 45(1):34–62, 2020. (Conference version in ICALP'12. All references are to the full preprint arXiv:1202.4798 .).

[Eve57] H. Everett. Recursive games. *Contributions to the Theory of Games*, 3(39):47–78, 1957.

[EWY19] K. Etessami, D. Wojtczak, and M. Yannakakis. Recursive Stochastic Games with Positive Rewards. *Theoretical Computer Science*, 777:308–328, 2019.

[EY06] K. Etessami and M. Yannakakis. Efficient Qualitative Analysis of classes of Recursive Markov Decision Processes and Simple Stochastic Games. In *Proc. of 23rd STACS*, 2006.

[EY08] K. Etessami and M. Yannakakis. Recursive Concurrent Stochastic Games. *Logical Methods in Computer Science (LMCS)*, 4(4), 2008.

[EY09] K. Etessami and M. Yannakakis. Recursive Markov Chains, Stochastic Grammars, and Monotone Systems of Nonlinear Equations. *Journal of the ACM*, 56(1), 2009.

[EY10] K. Etessami and M. Yannakakis. On the Complexity of Nash Equilibria and Other Fixed Points. *SIAM Journal on Computing*, 39(6):2531–2597, 2010.

[EY15] K. Etessami and M. Yannakakis. Recursive Markov Decision Processes and Recursive Stochastic Games. *Journal of the ACM*, 62(2):1–69, 2015.

[FM13] S. K. S Frederiksen and P. B. Miltersen. Approximating the Value of a Concurrent Reachability Game in the Polynomial Time Hierarchy. In *Proc. of 24th ISAAC*, pages 457–467, 2013.

[GGJ76] M. R. Garey, R. L. Graham, and D. S. Johnson. Some NP-complete geometric problems. In *Proc. of 8th STOC*, pages 10–22, 1976.

[Har63] T. E. Harris. *The Theory of Branching Processes*. Springer-Verlag, 1963.

[HIJM14] K. A. Hansen, R. Ibsen-Jensen, and P. B. Miltersen. The Complexity of Solving Reachability Games Using Value and Strategy Iteration. *Theory of Comp. Syst.*, 55(2):380–403, 2014.

[HJV05] P. Haccou, P. Jagers, and V. A. Vatutin. *Branching Processes: Variation, Growth, and Extinction of Populations*. Cambridge University Press, 2005.

[HK66] A. Hoffman and R. Karp. On Nonterminating Stochastic Games. *Management Science*, 12:359–370, 1966.

[HKM09] K. A. Hansen, M. Koucky, and P. B. Miltersen. Winning Concurrent Reachability Games Requires Doubly-Exponential Patience. In *Proc. of 24th LICS*, pages 332–341, 2009.

[KA02] M. Kimmel and D. E. Axelrod. *Branching processes in biology*. Springer, 2002.

[Kar84] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4(4):373–395, 1984.

[Kha79] L. G. Khachiyan. A Polynomial Algorithm in Linear Programming. *Soviet Math. Dokl.*, 20:191–194, 1979.

[LDK95] M. L. Littman, T. L. Dean, and L. P. Kaelbling. On the Complexity of Solving Markov Decision Problems. In *Proc. of the 11th conference on Uncertainty in artificial intelligence (UAI)*, pages 394–402, 1995.

[Mal99] G. Malajovich. An Effective Version of Kronecker's Theorem on Simultaneous Diophantine Approximation. Technical report, Department of Mathematics, City University of Hong Kong, Hong Kong, 1999.

[MM15] H. Michalewski and M. Mio. On the Problem of Computing the Probability of Regular Sets of Trees. In *Proc. of 35th FSTTCS*, pages 489–502, 2015.

[Pap94] C. Papadimitriou. On the Complexity of the Parity Argument and Other Inefficient Proofs of Existence. *J. of Computer and System Sciences*, 48(3):498–532, 1994.

[Pli76] S. R. Pliska. Optimization of Multitype Branching Processes. *Management Science*, 23(2):117–124, 1976.

[PS16] M. Przybyłko and M. Skrzypczak. On the Complexity of Branching Games with Regular Conditions. In *Proc. of 41st MFCS*, pages 1–14, 2016.

[Put94] M. L. Puterman. *Markov Decision Processes.* J. Wiley and Sons, 1994.

[RBCN13] G. Reiter, I. Bozic, K. Chatterjee, and M. A. Nowak. TTP: Tool for tumor progression. In *Proc. of 25th CAV*, pages 101–106, 2013.

[Ren92] J. Renegar. On the computational complexity and geometry of the first-order theory of the reals. parts I-III. *J. of Symbolic Computation*, 13(3):255–352, 1992.

[RW82] U. Rothblum and P. Whittle. Growth Optimality for Branching Markov Decision Chains. *Mathematics of Operations Research*, 7(4):582–601, 1982.

[SC85] A. P. Sistla and E. M. Clarke. The Complexity of Propositional Linear Temporal Logics. *J. of ACM*, 32(3):733–749, 1985.

[Sha53] S. Shapley. Stochastic Games. *Proc. of the National Academy of Sciences, USA*, 39:1095–1100, 1953.

[Som05] R. Somla. New Algorithms for Solving Simple Stochastic Games. *Electronic Notes in TCS*, 119(1):51–65, 2005.

[Tar55] A. Tarski. A lattice-theoretical fixpoint theorem and its applications. *Pacific Journal of Mathematics*, 5(2):285–309, 1955.

[Woj13] D. Wojtczak. Expected Termination Time in BPA Games. In *Proc. of 11th ATVA*, pages 303–318, 2013.

[Yan90] M. Yannakakis. The Analysis of Local Search Problems and Their Heuristics. In *Proc. of 7th STACS*, pages 298–311, 1990.