A Cancellation Law for Probabilistic Processes

A Cancellation Law for Probabilistic Processes

Rob van Glabbeek[*]
University of Edinburgh
University of New South Wales
rvg@stanford.edu

Jan Friso Groote
Eindhoven University of Technology
j.f.groote@tue.nl

Erik de Vink
Eindhoven University of Technology
evink@win.tue.nl

We show a cancellation property for probabilistic choice. If $\mu \oplus \rho$ and $\nu \oplus \rho$ are branching probabilistic bisimilar, then μ and ν are also branching probabilistic bisimilar. We do this in the setting of a basic process language involving non-deterministic and probabilistic choice and define branching probabilistic bisimilarity on distributions. Despite the fact that the cancellation property is very elegant and concise, we failed to provide a short and natural combinatorial proof. Instead we provide a proof using metric topology. Our major lemma is that every distribution can be unfolded into an equivalent stable distribution, where the topological arguments are required to deal with uncountable branching.

1 Introduction

A familiar property of the real numbers $\mathbb{R}$ is the additive cancellation law: if $x+z = y+z$ then $x = y$. Switching to the Boolean setting, and interpreting $+$ by $\vee$ and $=$ by $\Leftrightarrow$, the property becomes $(x \vee z) \Leftrightarrow (y \vee z)$ implies $x \Leftrightarrow y$. This is not generally valid. Namely, if z is true, nothing can be derived regarding the truth values of x and y. Algebraically speaking, the reals provide an 'additive inverse', and the Booleans do not have a 'disjunctive' version of it.

A similar situation holds for strong bisimilarity in the pure non-deterministic setting vs. strong bisimilarity in the mixed non-deterministic and probabilistic setting. When we have $E+G \leftrightarrow F+G$ for the non-deterministic processes $E+G$ and $F+G$, it may or may not be the case that $E \leftrightarrow F$. However, if $P\,{}_{1/2}{\oplus}\, R \leftrightarrow Q\,{}_{1/2}{\oplus}\, R$ for the probabilistic processes $P\,{}_{1/2}{\oplus}\, R$ and $Q\,{}_{1/2}{\oplus}\, R$, with probabilistic choice ${}_{1/2}{\oplus}$, we can exploit a semantic characterization of bisimilarity as starting point of a calculation. The characterization reads

$$P \leftrightarrow Q \quad \text{iff} \quad \forall C \in \mathscr{E}/{\leftrightarrow}\colon\, \mu[C] = \nu[C] \tag{1}$$

where the distributions $\mu, \nu \in Distr(\mathscr{E})$ are induced by P and Q. To spell out the above, two probabilistic processes P and Q are strongly bisimilar iff the distributions μ and ν induced by P and Q, respectively, assign the same probability to every equivalence class C of non-deterministic processes modulo strong bisimilarity. In the situation that $P\,{}_{1/2}{\oplus}\, R \leftrightarrow Q\,{}_{1/2}{\oplus}\, R$ we obtain from (1), for equivalence classes $C \in \mathscr{E}/{\leftrightarrow}$ and distributions μ, ν, and ρ induced by the processes P, Q, and R, that

$$\begin{aligned} P\,{}_{1/2}{\oplus}\, R \leftrightarrow Q\,{}_{1/2}{\oplus}\, R \implies \forall C \in \mathscr{E}/{\leftrightarrow}\colon\, \tfrac{1}{2}\mu[C] + \tfrac{1}{2}\rho[C] = \tfrac{1}{2}\nu[C] + \tfrac{1}{2}\rho[C] \implies \\ \forall C \in \mathscr{E}/{\leftrightarrow}\colon\, \tfrac{1}{2}\mu[C] = \tfrac{1}{2}\nu[C] \implies \forall C \in \mathscr{E}/{\leftrightarrow}\colon\, \mu[C] = \nu[C] \implies P \leftrightarrow Q \end{aligned}$$

relying on the arithmetic of the reals.

We are interested in whether the cancellation law also holds for weaker notions of process equivalence for probabilistic processes, especially for branching probabilistic bisimilarity as proposed in [16].

[*]Supported by Royal Society Wolfson Fellowship RSWF\R1\221008

G. Caltais and C. A. Mezzina (Eds): Combined Workshop on Expressiveness in Concurrency and Structural Operational Semantics (EXPRESS/SOS 2023).
EPTCS 387, 2023, pp. 42–58, doi:10.4204/EPTCS.387.5

We find that it does but the proof is involved. A number of initial attempts were directed towards finding a straightforward combinatorial proof, but all failed. A proof in a topological setting, employing the notion of sequential compactness to deal with potentially infinite sequences of transitions is reported in this paper. We leave the existence of a shorter, combinatorial proof as an open question.

Our strategy to prove the above cancellation law for probabilistic processes and branching probabilistic bisimilarity is based on two intermediate results: (i) every probabilistic process unfolds into a so-called *stable* probabilistic process, and (ii) for stable probabilistic processes a characterization of the form (1) does hold. Intuitively, a stable process is a process that cannot do an internal move without leaving its equivalence class.

In order to make the above more concrete, let us consider an example. For the ease of presentation we use distributions directly, rather than probabilistic processes. Let the distributions μ and ν be given by

$$\begin{aligned}\mu &= \tfrac{1}{2}\delta(a\cdot\partial(\mathbf{0})) \oplus \tfrac{1}{2}\delta(b\cdot\partial(\mathbf{0}))\\ \nu &= \tfrac{1}{3}\delta(\tau\cdot(\partial(a\cdot\partial(\mathbf{0}))\ {}_{\frac{1}{2}}\!\oplus\ \partial(b\cdot\partial(\mathbf{0})))) \oplus \tfrac{1}{3}\delta(a\cdot\partial(\mathbf{0})) \oplus \tfrac{1}{3}\delta(b\cdot\partial(\mathbf{0}))\end{aligned}$$

with a and b two different actions. The distribution μ assigns probability 0.5 to $a\cdot\partial(\mathbf{0})$, meaning an a-action followed by a deadlock with probability 1, and probability 0.5 to $b\cdot\partial(\mathbf{0})$, i.e. a b-action followed by deadlock with probability 1. The distribution ν assigns both these non-deterministic processes probability $\frac{1}{3}$ and assigns the remaining probability $\frac{1}{3}$ to $\tau\cdot(\partial(a\cdot\partial(\mathbf{0}))\ {}_{\frac{1}{2}}\!\oplus\ \partial(b\cdot\partial(\mathbf{0})))$, where a τ-action precedes a 50-50 percent choice between the processes mentioned earlier. Below, we show that μ and ν are branching probabilistic bisimilar, i.e. $\mu \leftrightarrow_b \nu$. However, if C_1, C_2 and C_3 are the three different equivalence classes of $\tau\cdot(\partial(a\cdot\partial(\mathbf{0}))\ {}_{\frac{1}{2}}\!\oplus\ \partial(b\cdot\partial(\mathbf{0})))$, $a\cdot\partial(\mathbf{0})$ and $b\cdot\partial(\mathbf{0})$, respectively, we have

$$\mu[C_1] = 0 \neq \tfrac{1}{3} = \nu[C_1],\ \mu[C_2] = \tfrac{1}{2} \neq \tfrac{1}{3} = \nu[C_2],\ \text{and}\ \mu[C_3] = \tfrac{1}{2} \neq \tfrac{1}{3} = \nu[C_3].$$

Thus, although $\mu \leftrightarrow_b \nu$, it does not hold that $\mu[C] = \nu[C]$ for every equivalence class C. Note that the distribution ν is not stable, in the sense that it allows an internal transition to the branchingly equivalent ν.

As indicated, we establish in this paper a cancellation law for branching probabilistic bisimilarity in the context of mixed non-deterministic and probabilistic choice, exploiting the process language of [6], while dealing with distributions of finite support over non-deterministic processes for its semantics. We propose the notion of a stable distribution and show that every distribution can be unfolded into a stable distribution by chasing its (partial) τ-transitions. Our framework, including the notion of branching probabilistic bisimulation, builds on that of [19, 16].

Another trait of the current paper, as in [19, 16], is that distributions are taken as semantic foundation for bisimilarity, rather than seeing bisimilarity primarily as an equivalence relation on non-deterministic processes, which is subsequently lifted to an equivalence relation on distributions, as is the case for the notion of branching probabilistic bisimilarity of [27, 26] and also of [2, 1]. The idea to consider distributions as first-class citizens for probabilistic bisimilarity stems from [11]. In the systematic overview of the spectrum [3], also Baier et al. argue that a behavioral relation on distributions is needed to properly deal with silent moves.

Metric spaces and complete metric spaces, as well as their associated categories, have various uses in concurrency theory. In the setting of semantics of probabilistic systems, metric topology has been advocated as underlying denotational domain, for example in [5, 21, 25]. For quantitative comparison of Markov systems, metrics and pseudo-metric have been proposed for a quantitative notion of behavior equivalence, see e.g. [10, 13, 7]. The specific use of metric topology in this paper to derive an existential property of a transition system seems new.

The remainder of the paper is organized as follows. Section 2 collects some definitions from metric topology and establishes some auxiliary results. A simple process language with non-deterministic and probabilistic choice is introduced in Section 3, together with examples and basic properties of the operational semantics. Our definition of branching probabilistic bisimilarity is given in Section 4, followed by a congruence result with respect to probabilistic composition and a confluence property. The main contribution of the paper is presented in Sections 5 and 6. Section 5 shows in a series of continuity lemmas that the set of branching probabilistic bisimilar descendants is a (sequentially) compact set. Section 6 exploits these results to argue that unfolding of a distribution by inert τ-transitions has a stable end point, meaning that a stable branchingly equivalent distribution can be reached. With that result in place, a cancellation law for branching probabilistic bisimilarity is established. Finally, Section 7 wraps up with concluding remarks and a discussion of future work.

2 Preliminaries

For a non-empty set X, we define $Distr(X)$ as the set of all probability distributions over X of finite support, i.e., $Distr(X) = \{\, \mu\colon X \to [0,1] \mid \sum_{x \in X} \mu(x) = 1,\, \mu(x) > 0$ for finitely many $x \in X \,\}$. We use $spt(\mu)$ to denote the finite set $\{\, x \in X \mid \mu(x) > 0 \,\}$. Often, we write $\mu = \bigoplus_{i \in I} p_i \cdot x_i$ for an index set I, $p_i \geqslant 0$ and $x_i \in X$ for $i \in I$, where $p_i > 0$ for finitely many $i \in I$. Implicitly, we assume $\sum_{i \in I} p_i = 1$. We also write $r\mu \oplus (1-r)\nu$ and, equivalently, $\mu \;{}_r{\oplus}\; \nu$ for $\mu, \nu \in Distr(X)$ and $0 \leqslant r \leqslant 1$. As expected, we have that $(r\mu \oplus (1-r)\nu)(x) = (\mu \;{}_r{\oplus}\; \nu)(x) = r\mu(x) + (1-r)\nu(x)$ for $x \in X$. The *Dirac distribution* on x, the unique distribution with support x, is denoted $\delta(x)$.

The set $Distr(X)$ becomes a complete[1] metric space when endowed with the sup-norm [14], given by $d(\mu, \nu) = \sup_{x \in X} |\mu(x) - \nu(x)|$. This distance is also known as the distance of uniform convergence or Chebyshev distance.

Theorem 1. *If $Y \subseteq X$ is finite, then $Distr(Y)$ is a sequentially compact subspace of $Distr(X)$. This means that every sequence in $Distr(Y)$ has a convergent subsequence with a limit in $Distr(Y)$.*

Proof. $Distr(Y)$ is a bounded subset of $\mathbb{R}^n$, where $n := |Y|$ is the size of Y. It also is closed. For $\mathbb{R}^n$ equipped with the Euclidean metric, the sequential compactness of closed and bounded subsets is known as the Bolzano-Weierstrass theorem [23]. When using the Chebyshev metric, the same proof applies. □

In Section 5 we use the topological structure of the set of distributions over non-deterministic processes to study unfolding of partial τ-transitions. There we make use of the following representation property.

Lemma 2. *Suppose the sequence of distributions $(\mu_i)_{i=0}^{\infty}$ converges to the distribution μ in $Distr(X)$. Then a sequence of distributions $(\mu_i')_{i=0}^{\infty}$ in $Distr(X)$ and a sequence of probabilities $(r_i)_{i=0}^{\infty}$ in $[0,1]$ exist such that $\mu_i = (1-r_i)\mu \oplus r_i\mu_i'$ for $i \in \mathbb{N}$ and $\lim_{i \to \infty} r_i = 0$.*

Proof. Let $i \in \mathbb{N}$. For $x \in spt(\mu)$, the quotient $\mu_i(x)/\mu(x)$ is non-negative, but may exceed 1. However, $0 \leqslant \min\{ \frac{\mu_i(x)}{\mu(x)} \mid x \in spt(\mu) \} \leqslant 1$, since the numerator cannot strictly exceed the denominator for all $x \in spt(\mu)$. Let $r_i = 1 - \min\{ \frac{\mu_i(x)}{\mu(x)} \mid x \in spt(\mu) \}$ for $i \in \mathbb{N}$. Then we have $0 \leqslant r_i \leqslant 1$.

[1] A *Cauchy sequence* is a sequence of points in a metric space whose elements become arbitrarily close to each other as the sequence progresses. The space is *complete* if every such sequence has a limit within the space.

For $i \in \mathbb{N}$, define $\mu_i' \in Distr(X)$ as follows. If $r_i > 0$ then $\mu_i'(x) = 1/r_i \cdot \big[\mu_i(x) - (1-r_i)\mu(x)\big]$ for $x \in X$; if $r_i = 0$ then $\mu_i' = \mu$. We verify for $r_i > 0$ that μ_i' is indeed a distribution: (i) For $x \notin spt(\mu)$ it holds that $\mu(x) = 0$, and therefore $\mu_i'(x) = 1/r_i \cdot \mu_i(x) \geqslant 0$. For $x \in spt(\mu)$,

$$\mu_i'(x) = 1/r_i \cdot \big[\mu_i(x) - (1-r_i)\mu(x)\big] = \mu(x)/r_i \cdot \big[\frac{\mu_i(x)}{\mu(x)} - \frac{\mu_i(x_{min})}{\mu(x_{min})}\big] \geqslant 0$$

for $x_{min} \in spt(\mu)$ such that $\mu_i(x_{min})/\mu(x_{min})$ is minimal. (ii) In addition,

$$\begin{aligned}\textstyle\sum\{\, \mu_i'(x) \mid x \in X \,\} &= 1/r_i \cdot \textstyle\sum\{\, \mu_i(x) \mid x \notin spt(\mu) \,\} + 1/r_i \cdot \textstyle\sum\{\, \mu_i(x) - (1-r_i)\mu(x) \mid x \in spt(\mu) \,\} = \\ &1/r_i \cdot \textstyle\sum\{\, \mu_i(x) \mid x \in X \,\} - (1-r_i)/r_i \cdot \textstyle\sum\{\, \mu(x) \mid x \in spt(\mu) \,\} = 1/r_i - (1-r_i)/r_i = r_i/r_i = 1.\end{aligned}$$

Therefore, $0 \leqslant \mu_i'(x) \leqslant 1$ and $\sum\{\, \mu_i'(x) \mid x \in X \,\} = 1$.

Now we prove that $\mu_i = (1-r_i)\mu \oplus r_i\mu_i'$. If $r_i = 0$, then $\mu_i = \mu$, $\mu_i' = \mu$, and $\mu_i = (1-r_i)\mu \oplus r_i\mu_i'$. If $r_i > 0$, then $\mu_i(x) = (1-r_i)\mu(x) \oplus r_i\mu_i'(x)$ by definition of $\mu_i'(x)$ for all $x \in X$. Thus, also $\mu_i = (1-r_i)\mu \oplus r_i\mu_i'$ in this case.

Finally, we show that $\lim_{i\to\infty} r_i = 0$. Let $x'_{min} \in spt(\mu)$ be such that $\mu(x'_{min})$ is minimal. Then we have

$$r_i = 1 - \min\{\, \frac{\mu_i(x)}{\mu(x)} \mid x \in spt(\mu) \,\} = \max\{\, \frac{\mu(x) - \mu_i(x)}{\mu(x)} \mid x \in spt(\mu),\, \mu(x) \geqslant \mu_i(x) \,\} \leqslant \frac{d(\mu,\mu_i)}{\mu(x'_{min})}$$

By assumption, $\lim_{i\to\infty} d(\mu,\mu_i) = 0$. Hence also $\lim_{i\to\infty} r_i = 0$, as was to be shown. □

The following combinatorial result is helpful in the sequel.

Lemma 3. *Let I and J be finite index sets, $p_i, q_j \in [0,1]$ and $\mu_i, \nu_j \in Distr(X)$, for $i \in I$ and $j \in J$, such that $\bigoplus_{i\in I} p_i\mu_i = \bigoplus_{j\in J} q_j\nu_j$. Then $r_{ij} \geqslant 0$ and $\rho_{ij} \in Distr(X)$ exist such that $\sum_{j\in J} r_{ij} = p_i$ and $p_i \cdot \mu_i = \bigoplus_{j\in J} r_{ij} \cdot \rho_{ij}$ for all $i \in I$, and $\sum_{i\in I} r_{ij} = q_j$ and $q_j \cdot \nu_j = \bigoplus_{i\in I} r_{ij} \cdot \rho_{ij}$ for all $j \in J$.*

Proof. Let $\xi = \bigoplus_{i\in I} p_i \cdot \mu_i = \bigoplus_{j\in J} q_j \cdot \nu_j$. We define $r_{ij} = \sum_{x\in spt(\xi)} \dfrac{p_i\mu_i(x) \cdot q_j\nu_j(x)}{\xi(x)}$ for all $i \in I$ and $j \in J$. In case $r_{ij} = 0$, choose $\rho_{ij} \in Distr(X)$ arbitrarily. In case $r_{ij} \neq 0$, define $\rho_{ij} \in Distr(X)$, for $i \in I$ and $j \in J$, by

$$\rho_{ij}(x) = \begin{cases} \dfrac{p_i\mu_i(x) \cdot q_j\nu_j(x)}{r_{ij}\,\xi(x)} & \text{if } \xi(x) > 0, \\ 0 & \text{otherwise} \end{cases}$$

for all $x \in X$. By definition of r_{ij} and ρ_{ij} it holds that $\sum\{\, \rho_{ij}(x) \mid x \in X \,\} = 1$. So, $\rho_{ij} \in Distr(X)$ indeed.

We verify $\sum_{j\in J} r_{ij} = p_i$ and $p_i \cdot \mu_i = \bigoplus_{j\in J} r_{ij} \cdot \rho_{ij}$ for $i \in I$.

$$\begin{aligned}\textstyle\sum_{j\in J} r_{ij} &= \textstyle\sum_{j\in J} \sum_{x\in spt(\xi)} p_i\mu_i(x) \cdot q_j\nu_j(x)/\xi(x) \\ &= \textstyle\sum_{x\in spt(\xi)} p_i\mu_i(x) \cdot \sum_{j\in J} q_j\nu_j(x)/\xi(x) \\ &= \textstyle\sum_{x\in spt(\xi)} p_i\mu_i(x) && \text{(since } \xi = \textstyle\bigoplus_{j\in J} q_j \cdot \nu_j) \\ &= p_i \textstyle\sum_{x\in spt(\xi)} \mu_i(x) \\ &= p_i.\end{aligned}$$

Next, pick $y \in X$ and $i \in I$. If $\xi(y) = 0$, then $p_i\mu_i(y) = 0$, since $\xi(y) = \sum_{i\in I} p_i\mu_i(y)$, and $r_{ij} = 0$ or $\rho_{ij}(y) = 0$ for all $j \in J$, by the various definitions, thus $\sum_{j\in J} r_{ij}\rho_{ij}(y) = 0$ as well.

Suppose $\xi(y) > 0$. Put $J_i = \{ j \in J \mid r_{ij} > 0 \}$. If $j \in J \backslash J_i$, i.e. if $r_{ij} = 0$, then $p_i \mu_i(y) q_j \nu_j(y)/\xi(y) = 0$ by definition of r_{ij}. Therefore we have

$$\begin{aligned}
\textstyle\sum_{j \in J} r_{ij} \rho_{ij}(y) &= \textstyle\sum_{j \in J_i} r_{ij} \rho_{ij}(y) \\
&= \textstyle\sum_{j \in J_i} r_{ij} p_i \mu_i(y) \cdot q_j \nu_j(y) / (r_{ij} \xi(y)) \\
&= \textstyle\sum_{j \in J_i} p_i \mu_i(y) \cdot q_j \nu_j(y) / \xi(y) \\
&= \textstyle\sum_{j \in J} p_i \mu_i(y) \cdot q_j \nu_j(y) / \xi(y) && \text{(summand zero for } j \in J \backslash J_i) \\
&= p_i \mu_i(y) / \xi(y) \cdot \textstyle\sum_{j \in J} q_j \nu_j(y) \\
&= p_i \mu_i(y) && \text{(since } \xi = \textstyle\bigoplus_{j \in J} q_j \cdot \nu_j).
\end{aligned}$$

The statements $\sum_{i \in I} r_{ij} = q_j$ and $q_j \cdot \nu_j = \bigoplus_{i \in I} r_{ij} \cdot \rho_{ij}$ for $j \in J$ follow by symmetry. □

3 An elementary processes language

In this section we define a syntax and transition system semantics for non-deterministic and probabilistic processes. Depending on the top operator, following [6], a process is either a non-deterministic process $E \in \mathscr{E}$, with constant $\mathbf{0}$, prefix operators $\alpha\cdot$ and non-deterministic choice $+$, or a probabilistic process $P \in \mathscr{P}$, with the Dirac operator ∂ and probabilistic choices ${}_r\oplus$.

Definition 4 (Syntax)**.** *The classes $\mathscr{E}$ and $\mathscr{P}$ of non-deterministic and probabilistic processes, respectively, over the set of actions $\mathscr{A}$, are given by*

$$E ::= \mathbf{0} \mid \alpha \cdot P \mid E + E \qquad\qquad P ::= \partial(E) \mid P \,{}_r\!\oplus P$$

with actions α from $\mathscr{A}$ and where $0 \leqslant r \leqslant 1$.

We use $E, F, \dots$ to range over $\mathscr{E}$ and $P, Q, \dots$ to range over $\mathscr{P}$. The probabilistic process $P_1 \,{}_r\!\oplus P_2$ behaves as P_1 with probability r and behaves as P_2 with probability $1 - r$.

We introduce a complexity measure $c : \mathscr{E} \cup \mathscr{P} \to \mathbb{N}$ for non-deterministic and probabilistic processes based on the size of a process. It is given by $c(\mathbf{0}) = 0$, $c(a \cdot P) = c(P) + 1$, $c(E + F) = c(E) + c(F)$, and $c(\partial(E)) = c(E) + 1$, $c(P \,{}_r\!\oplus Q) = c(P) + c(Q)$.

Examples As illustration, we provide the following pairs of non-deterministic processes, which are branching probabilistic bisimilar in the sense of Definition 9.

(i) $\mathbf{H_1} = a \cdot \big(P \,{}_{\frac{1}{4}}\!\oplus (P \,{}_{\frac{1}{3}}\!\oplus Q)\big)$ and $\mathbf{H_2} = a \cdot \big(P \,{}_{\frac{1}{2}}\!\oplus (Q \,{}_{\frac{1}{2}}\!\oplus Q)\big)$

(ii) $\mathbf{G_1} = a \cdot (P \,{}_{\frac{1}{2}}\!\oplus Q)$ and $\mathbf{G_2} = a \cdot \big(\partial\big(\tau \cdot (P \,{}_{\frac{1}{2}}\!\oplus Q)\big) \,{}_{\frac{1}{3}}\!\oplus (P \,{}_{\frac{1}{2}}\!\oplus Q)\big)$

(iii) $\mathbf{I_1} = a \cdot \partial(b \cdot P + \tau \cdot Q)$ and $\mathbf{I_2} = a \cdot \partial(\tau \cdot \partial(b \cdot P + \tau \cdot Q) + b \cdot P + \tau \cdot Q)$

The examples $\mathbf{H_1}$ and $\mathbf{H_2}$ are taken from [22], and $\mathbf{G_1}$ and $\mathbf{G_2}$ are taken from [16]. The processes $\mathbf{G_2}$ and $\mathbf{I_2}$ contain a so-called inert τ-transition.

As usual, the SOS semantics for $\mathscr{E}$ and $\mathscr{P}$ makes use of two types of transition relations [20, 6, 16].

Definition 5 (Operational semantics).
(a) The transition relations $\rightarrow \subseteq \mathcal{E} \times \mathcal{A} \times Distr(\mathcal{E})$ *and* $\mapsto \subseteq \mathcal{P} \times Distr(\mathcal{E})$ *are given by*

$$\frac{P \mapsto \mu}{\alpha \cdot P \xrightarrow{\alpha} \mu}\ (\text{PREF})$$

$$\frac{E_1 \xrightarrow{\alpha} \mu_1}{E_1 + E_2 \xrightarrow{\alpha} \mu_1}\ (\text{ND-CHOICE 1}) \qquad \frac{E_2 \xrightarrow{\alpha} \mu_2}{E_1 + E_2 \xrightarrow{\alpha} \mu_2}\ (\text{ND-CHOICE 2})$$

$$\frac{}{\partial(E) \mapsto \delta(E)}\ (\text{DIRAC}) \qquad \frac{P_1 \mapsto \mu_1 \quad P_2 \mapsto \mu_2}{P_1\,{}_r\!\oplus P_2 \mapsto \mu_1\,{}_r\!\oplus \mu_2}\ (\text{P-CHOICE})$$

(b) The transition relation $\rightarrow \subseteq Distr(\mathcal{E}) \times \mathcal{A} \times Distr(\mathcal{E})$ *is such that* $\mu \xrightarrow{\alpha} \mu'$ *whenever* $\mu = \bigoplus_{i \in I} p_i \cdot E_i$, $\mu' = \bigoplus_{i \in I} p_i \cdot \mu_i'$, *and* $E_i \xrightarrow{\alpha} \mu_i'$ *for all* $i \in I$.

In rule (DIRAC) of the relation $\mapsto$ we have that the syntactic Dirac process $\partial(E)$ is coupled to the semantic Dirac distribution $\delta(E)$. Similarly, in (P-CHOICE), the syntactic probabilistic operator ${}_r\!\oplus$ in $P_1\,{}_r\!\oplus P_2$ is replaced by semantic probabilistic composition in $\mu_1\,{}_r\!\oplus \mu_2$. Thus, with each probabilistic process $P \in \mathcal{P}$ we associate a distribution $[\![P]\!] \in Distr(\mathcal{E})$ as follows: $[\![\partial(E)]\!] = \delta(E)$ and $[\![P\,{}_r\!\oplus Q]\!] = [\![P]\!]\,{}_r\!\oplus [\![Q]\!]$, which is the distribution $r[\![P]\!] \oplus (1-r)[\![Q]\!]$.

The relation $\rightarrow$ for non-deterministic processes is finitely branching, but the relation $\rightarrow$ for probabilistic processes is not. Following [27, 26], the transition relation $\rightarrow$ on distributions as given by Definition 5 allows for a probabilistic combination of non-deterministic alternatives resulting in a so-called combined transition. For example, for the process $E = a \cdot (P\,{}_{\frac{1}{2}}\!\oplus Q) + a \cdot (P\,{}_{\frac{1}{3}}\!\oplus Q)$ of [6], we have that the Dirac process $\partial(E) = \partial(a \cdot (P\,{}_{\frac{1}{2}}\!\oplus Q) + a \cdot (P\,{}_{\frac{1}{3}}\!\oplus Q))$ provides an a-transition to $[\![P\,{}_{\frac{1}{2}}\!\oplus Q]\!]$ as well as an a-transition to $[\![P\,{}_{\frac{1}{3}}\!\oplus Q]\!]$. So, since we can represent the distribution $\delta(E)$ by $\delta(E) = \frac{1}{2}\delta(E) \oplus \frac{1}{2}\delta(E)$, the distribution $\delta(E)$ also has a combined transition

$$\delta(E) = \tfrac{1}{2}\delta(E) \oplus \tfrac{1}{2}\delta(E) \xrightarrow{a} \tfrac{1}{2}[\![P\,{}_{\frac{1}{2}}\!\oplus Q]\!] \oplus \tfrac{1}{2}[\![P\,{}_{\frac{1}{3}}\!\oplus Q]\!] = [\![P\,{}_{\frac{5}{12}}\!\oplus Q]\!].$$

As noted in [28], the ability to combine transitions is crucial for obtaining transitivity of probabilistic process equivalences that take internal actions into account.

Example Referring to the examples of processes above, we have, e.g,

$$\begin{aligned}
\mathbf{H_1}: &\quad \delta(a \cdot (P\,{}_{\frac{1}{4}}\!\oplus (P\,{}_{\frac{1}{3}}\!\oplus Q))) \xrightarrow{a} [\![P\,{}_{\frac{1}{4}}\!\oplus (P\,{}_{\frac{1}{3}}\!\oplus Q)]\!] = \tfrac{1}{2}[\![P]\!] \oplus \tfrac{1}{2}[\![Q]\!] \\
\mathbf{H_2}: &\quad \delta(a \cdot (P\,{}_{\frac{1}{2}}\!\oplus (Q\,{}_{\frac{1}{2}}\!\oplus Q))) \xrightarrow{a} [\![P\,{}_{\frac{1}{2}}\!\oplus (Q\,{}_{\frac{1}{2}}\!\oplus Q)]\!] = \tfrac{1}{2}[\![P]\!] \oplus \tfrac{1}{2}[\![Q]\!] \\
\mathbf{G_2}: &\quad a \cdot \big(\partial\big(\tau \cdot (P\,{}_{\frac{1}{2}}\!\oplus Q)\big)\,{}_{\frac{1}{3}}\!\oplus (P\,{}_{\frac{1}{2}}\!\oplus Q)\big) \xrightarrow{a} \delta\big(\tau \cdot (P\,{}_{\frac{1}{2}}\!\oplus Q)\big)\,{}_{\frac{1}{3}}\!\oplus (P\,{}_{\frac{1}{2}}\!\oplus Q).
\end{aligned}$$

Because a transition of a probabilistic process yields a distribution, the a-transitions of $\mathbf{H_1}$ and $\mathbf{H_2}$ have the same target. It is noted that $\mathbf{G_2}$ doesn't provide a further transition unless both its components P and Q do so to match the transition of $\tau \cdot (P\,{}_{\frac{1}{2}}\!\oplus Q)$.

In preparation to the definition of the notion of branching probabilistic bisimilarity in Section 4 we introduce some notation.

Definition 6. *For* $\mu, \mu' \in Distr(\mathcal{E})$ *and* $\alpha \in \mathcal{A}$ *we write* $\mu \xrightarrow{(\alpha)} \mu'$ *iff (i)* $\mu \xrightarrow{\alpha} \mu'$, *or (ii)* $\alpha = \tau$ *and* $\mu' = \mu$, *or (iii)* $\alpha = \tau$ *and there exist* $\mu_1, \mu_2, \mu_1', \mu_2' \in Distr(\mathcal{E})$ *such that* $\mu = \mu_1\,{}_r\!\oplus \mu_2$, $\mu' = \mu_1'\,{}_r\!\oplus \mu_2'$, $\mu_1 \xrightarrow{\tau} \mu_1'$ *and* $\mu_2 = \mu_2'$ *for some* $r \in (0,1)$.

Cases (i) and (ii) in the definition above correspond with the limits $r = 1$ and $r = 0$ of case (iii). We use $\Rightarrow$ to denote the reflexive transitive closure of $\xrightarrow{(\tau)}$. A transition $\mu \xrightarrow{(\tau)} \mu'$ is called a partial transition, and a transition $\mu \Rightarrow \mu'$ is called a weak transition.

Example

(a) According to Definition 6 we have

$$\tfrac{1}{3}\delta(\tau\cdot(P\,{}_{\frac{1}{2}}{\oplus}\,Q))\oplus\tfrac{2}{3}[\![P\,{}_{\frac{1}{2}}{\oplus}\,Q]\!] \xrightarrow{(\tau)} \tfrac{1}{3}[\![P\,{}_{\frac{1}{2}}{\oplus}\,Q]\!]\oplus\tfrac{2}{3}[\![P\,{}_{\frac{1}{2}}{\oplus}\,Q]\!]=[\![P\,{}_{\frac{1}{2}}{\oplus}\,Q]\!].$$

(b) There are typically multiple ways to construct a weak transition $\Rightarrow$. Consider the weak transition $\frac{1}{2}\delta(\tau\cdot\partial(\tau\cdot P))\oplus\frac{1}{3}\delta(\tau\cdot P)\oplus\frac{1}{6}[\![P]\!]\Rightarrow[\![P]\!]$ which can be obtained, among uncountably many other possibilities, via

$$\begin{aligned}&\tfrac{1}{2}\delta(\tau\cdot\partial(\tau\cdot P))\oplus\tfrac{1}{3}\delta(\tau\cdot P)\oplus\tfrac{1}{6}[\![P]\!]\xrightarrow{(\tau)}\\&\qquad\tfrac{1}{2}\delta(\tau\cdot P))\oplus\tfrac{1}{3}\delta(\tau\cdot P)\oplus\tfrac{1}{6}[\![P]\!]=\tfrac{5}{6}\delta(\tau\cdot P)\oplus\tfrac{1}{6}[\![P]\!]\xrightarrow{(\tau)}[\![P]\!],\end{aligned}$$

or via

$$\begin{aligned}&\tfrac{1}{2}\delta(\tau\cdot\partial(\tau\cdot P))\oplus\tfrac{1}{3}\delta(\tau\cdot P)\oplus\tfrac{1}{6}[\![P]\!]\xrightarrow{(\tau)}\tfrac{1}{2}\delta(\tau\cdot\partial(\tau\cdot P))\oplus\tfrac{1}{3}\delta(P)\oplus\tfrac{1}{6}[\![P]\!]=\\&\qquad\tfrac{1}{2}\delta(\tau\cdot\partial(\tau\cdot P))\oplus\tfrac{1}{2}\delta(P)\xrightarrow{(\tau)}\tfrac{1}{2}\delta(\tau\cdot P)\oplus\tfrac{1}{2}[\![P]\!]\xrightarrow{(\tau)}\tfrac{1}{2}[\![P]\!]\oplus\tfrac{1}{2}[\![P]\!]=[\![P]\!].\end{aligned}$$

(c) The distribution $\frac{1}{2}\delta(\tau\cdot\partial(a\cdot\partial(\mathbf{0})+b\cdot\partial(\mathbf{0})))\oplus\frac{1}{2}\delta(a\cdot\partial(c\cdot\partial(\mathbf{0})))$ doesn't admit a τ-transition nor an a-transition. However, we have

$$\begin{aligned}&\tfrac{1}{2}\delta(\tau\cdot\partial(a\cdot\partial(\mathbf{0})+b\cdot\partial(\mathbf{0})))\oplus\tfrac{1}{2}\delta(a\cdot\partial(c\cdot\partial(\mathbf{0})))\xrightarrow{(\tau)}\\&\qquad\tfrac{1}{2}\partial(a\cdot\partial(\mathbf{0})+b\cdot\partial(\mathbf{0}))\oplus\tfrac{1}{2}\delta(a\cdot\partial(c\cdot\partial(\mathbf{0})))\xrightarrow{a}\tfrac{1}{2}\delta(\mathbf{0})\oplus\tfrac{1}{2}\delta(c\cdot\partial(\mathbf{0})).\end{aligned}$$

The following lemma states that the transitions $\xrightarrow{\alpha}$, $\xrightarrow{(\alpha)}$, and $\Rightarrow$ of Definitions 5 and 6 can be probabilistically composed.

Lemma 7. *Let, for a finite index set I, $\mu_i,\mu_i'\in Distr(\mathscr{E})$ and $p_i\geqslant 0$ such that $\sum_{i\in I}p_i=1$.*

(a) If $\mu_i\xrightarrow{\alpha}\mu_i'$ for all $i\in I$, then $\bigoplus_{i\in I}p_i\cdot\mu_i\xrightarrow{\alpha}\bigoplus_{i\in I}p_i\cdot\mu_i'$.

(b) If $\mu_i\xrightarrow{(\tau)}\mu_i'$ for all $i\in I$, then $\bigoplus_{i\in I}p_i\cdot\mu_i\xrightarrow{(\tau)}\bigoplus_{i\in I}p_i\cdot\mu_i'$.

(c) If $\mu_i\Rightarrow\mu_i'$ for all $i\in I$, then $\bigoplus_{i\in I}p_i\cdot\mu_i\Rightarrow\bigoplus_{i\in I}p_i\cdot\mu_i'$.

Proof. Let $\mu=\bigoplus_{i\in I}p_i\cdot\mu_i$ and $\mu'=\bigoplus_{i\in I}p_i\cdot\mu_i'$. Without loss of generality, we may assume that $p_i>0$ for all $i\in I$.

(a) Suppose $\mu_i\xrightarrow{\alpha}\mu_i'$ for all $i\in I$. Then, by Definition 5, $\mu_i=\bigoplus_{j\in J_i}p_{ij}\cdot E_{ij}$, $\mu_i'=\bigoplus_{j\in J_i}p_{ij}\cdot\eta_{ij}$, and $E_{ij}\xrightarrow{\alpha}\eta_{ij}$ for $j\in J_i$ for a suitable index set J_i, $p_{ij}>0$ and $\eta_{ij}\in Distr(\mathscr{E})$. Define the index set K and probabilities q_k for $k\in K$ by $K=\{\,(i,j)\mid i\in I,\,j\in J_i\,\}$ and $q_{(i,j)}=p_ip_{ij}$ for $(i,j)\in K$, so that $\sum_{k\in K}q_k=1$. Then we have $\mu=\bigoplus_{k\in K}q_k\cdot E_{ij}$ and $\mu'=\bigoplus_{k\in K}q_k\cdot\eta_{ij}$. Therefore, by Definition 5, it follows that $\mu\xrightarrow{\alpha}\mu'$.

(b) Let $\mu_i\xrightarrow{(\tau)}\mu_i'$ for all $i\in I$. Then, for all $i\in I$, by Definition 6, there exists $r_i\in[0,1]$ and $\mu_i^{\text{stay}},\mu_i^{\text{go}},\mu_i''\in Distr(\mathscr{E})$, such that $\mu_i=\mu_i^{\text{stay}}\,{}_{r_i}{\oplus}\,\mu_i^{\text{go}}$, $\mu_i'=\mu_i^{\text{stay}}\,{}_{r_i}{\oplus}\,\mu_i''$, and either $r_i=1$ or $\mu_i^{\text{go}}\xrightarrow{\tau}\mu_i''$. In case $r_i=0$ for all $i\in I$, we have that $\mu_i\xrightarrow{\tau}\mu_i'$ for all $i\in I$, and thus $\mu\xrightarrow{\tau}\mu'$ by the first claim of the lemma, and $\mu\xrightarrow{(\tau)}\mu'$ by Definition 6(i). In case $r_i=1$ for all $i\in I$, we have $\mu'=\mu$ and thus $\mu\xrightarrow{(\tau)}\mu'$ by Definition 6(ii). Otherwise, let $I':=\{\,i\in I\mid r_i<1\,\}$, $r=\sum_{i\in I}p_i\cdot r_i$, $\mu^{\text{stay}}:=\bigoplus_{i\in I}\frac{p_i\cdot r_i}{r}\cdot\mu_i^{\text{stay}}$, $\mu^{\text{go}}:=\bigoplus_{i\in I'}\frac{p_i\cdot(1-r_i)}{1-r}\cdot\mu_i^{\text{go}}$ and $\mu'':=\bigoplus_{i\in I'}\frac{p_i\cdot(1-r_i)}{1-r}\cdot\mu_i''$. Then $\mu^{\text{go}}\xrightarrow{\tau}\mu''$ by the first claim of the lemma. Moreover, $\mu=\mu^{\text{stay}}\,{}_r{\oplus}\,\mu^{\text{go}}$, $\mu'=\mu^{\text{stay}}\,{}_r{\oplus}\,\mu''$ and $r\in(0,1)$. So $\mu\xrightarrow{(\tau)}\mu'$ by Definition 6(iii).

(c) Let $\mu_i\Rightarrow\mu_i'$ for all $i\in I$. As I is finite and $\Rightarrow$ is reflexive, there exists an $n\in\mathbb{N}$ such that $\mu_i=\mu_i^{(0)}\xrightarrow{(\tau)}\mu_i^{(1)}\xrightarrow{(\tau)}\dots\xrightarrow{(\tau)}\mu_i^{(n)}=\mu_i'$ for all $i\in I$. Now $\mu\Rightarrow\mu'$ follows by n applications of the second statement of the lemma. $\square$

Likewise, the next lemma allows *probabilistic decomposition* of transitions $\xrightarrow{\alpha}$, $\xrightarrow{(\alpha)}$ and $\Rightarrow$.

Lemma 8. *Let $\mu, \mu' \in Distr(\mathscr{E})$ and $\mu = \bigoplus_{i\in I} p_i \cdot \mu_i$ with $p_i > 0$ for $i \in I$.*

(a) If $\mu \xrightarrow{\alpha} \mu'$, then there are μ_i' for $i \in I$ such that $\mu_i \xrightarrow{\alpha} \mu_i'$ for $i \in I$ and $\mu' = \bigoplus_{i\in I} p_i \cdot \mu_i'$.

(b) If $\mu \xrightarrow{(\tau)} \mu'$, then there are μ_i' for $i \in I$ such that $\mu_i \xrightarrow{(\tau)} \mu_i'$ for $i \in I$ and $\mu' = \bigoplus_{i\in I} p_i \cdot \mu_i'$.

(c) If $\mu \Rightarrow \mu'$, then there are μ_i' for $i \in I$ such that $\mu_i \Rightarrow \mu_i'$ for $i \in I$ and $\mu' = \bigoplus_{i\in I} p_i \cdot \mu_i'$.

Proof. (a) Suppose $\mu \xrightarrow{\alpha} \mu'$. By Definition 5 $\mu = \bigoplus_{j\in J} q_j \cdot E_j$, $\mu' = \bigoplus_{j\in J} q_j \cdot \eta_j$, and $E_j \xrightarrow{\alpha} \eta_j$ for all $j \in J$, for suitable index set J, $q_j > 0$, $E_j \in \mathscr{E}$, and $\eta_j \in Distr(\mathscr{E})$. By Lemma 3 there are $r_{ij} \geqslant 0$ and $\rho_{ij} \in Distr(\mathscr{E})$ such that $\sum_{j\in J} r_{ij} = p_i$ and $p_i\mu_i = \bigoplus_{j\in J} r_{ij}\rho_{ij}$ for $i \in I$, and $\sum_{i\in I} r_{ij} = q_j$ and $q_j \cdot \delta(E_j) = \bigoplus_{i\in I} r_{ij}\rho_{ij}$ for all $j \in J$. Hence, $\rho_{ij} = \delta(E_j)$ for $i \in I$, $j \in J$.

For all $i \in I$, let $\mu_i' = \bigoplus_{j\in J} (r_{ij}/p_i)\eta_j$. Then $\mu_i \xrightarrow{\alpha} \mu_i'$, for all $i \in I$, by Lemma 7(a). Moreover, it holds that $\bigoplus_{i\in I} p_i\mu_i' = \bigoplus_{i\in I} p_i \cdot \bigoplus_{j\in J} (r_{ij}/p_i)\eta_j = \bigoplus_{j\in J} \bigoplus_{i\in I} r_{ij} \cdot \eta_j = \bigoplus_{j\in J} q_j \cdot \eta_j = \mu'$.

(b) Suppose $\mu \xrightarrow{(\tau)} \mu'$. By Definition 6, either (i) $\mu \xrightarrow{\tau} \mu'$, or (ii) $\mu' = \mu$, or (iii) there exist $\nu_1, \nu_2, \nu_1', \nu_2' \in Distr(\mathscr{E})$ such that $\mu = \nu_1 \,{}_r\!\oplus\, \nu_2$, $\mu' = \nu_1' \,{}_r\!\oplus\, \nu_2'$, $\nu_1 \xrightarrow{\tau} \nu_1'$ and $\nu_2 = \nu_2'$ for some $r \in (0,1)$. In case (i), the required μ_i' exist by the first statement of this lemma. In case (ii) one can simply take $\mu_i' := \mu_i$ for all $i \in I$. Hence assume that case (iii) applies. Let $J := \{1,2\}$, $q_1 := r$ and $q_2 := 1-r$. By Lemma 3 there are $r_{ij} \in [0,1]$ and $\rho_{ij} \in Distr(\mathscr{E})$ with $\sum_{j\in J} r_{ij} = p_i$ and $\mu_i = \bigoplus_{j\in J} \frac{r_{ij}}{p_i} \cdot \rho_{ij}$ for all $i \in I$, and $\sum_{i\in I} r_{ij} = q_j$ and $\nu_j = \bigoplus_{i\in I} \frac{r_{ij}}{q_j} \cdot \rho_{ij}$ for all $j \in J$.

Let $I' := \{i \in I \mid r_{i1} > 0\}$. Since $\nu_1 = \bigoplus_{i\in I'} \frac{r_{i1}}{r} \cdot \rho_{i1} \xrightarrow{\tau} \nu_1'$, by the first statement of the lemma, for all $i \in I'$ there are ρ_{i1}' such that $\rho_{i1} \xrightarrow{\tau} \rho_{i1}'$ and $\nu_1' = \bigoplus_{i\in I'} \frac{r_{i1}}{r} \cdot \rho_{i1}'$. For all $i \in I \backslash I'$ pick $\rho_{i1}' \in Distr(\mathscr{E})$ arbitrarily. It follows that $\mu_i = \rho_{i1} \,{}_{\frac{r_{i1}}{p_i}}\!\oplus\, \rho_{i2} \xrightarrow{(\tau)} \rho_{i1}' \,{}_{\frac{r_{i1}}{p_i}}\!\oplus\, \rho_{i2} =: \mu_i'$ for all $i \in I$. Moreover,

$$\bigoplus_{i\in I} p_i \cdot \mu_i' = \bigoplus_{i\in I} p_i \cdot (\rho_{i1}' \,{}_{\frac{r_{i1}}{p_i}}\!\oplus\, \rho_{i2}) = (\bigoplus_{i\in I} \tfrac{r_{i1}}{r} \cdot \rho_{i1}') \,{}_r\!\oplus\, (\bigoplus_{i\in I} \tfrac{r_{i2}}{1-r} \cdot \rho_{i2}) = \nu_1' \,{}_r\!\oplus\, \nu_2 = \mu' .$$

(c) The last statement follows by transitivity from the second one. □

4 Branching probabilistic bisimilarity

In this section we recall the notion of branching probabilistic bisimilarity [16]. The notion is based on a decomposability property due to [9] and a transfer property.

Definition 9 (Branching probabilistic bisimilarity)**.**

(a) A relation $\mathscr{R} \subseteq Distr(\mathscr{E}) \times Distr(\mathscr{E})$ is called weakly decomposable iff it is symmetric and for all $\mu, \nu \in Distr(\mathscr{E})$ such that $\mu \mathscr{R} \nu$ and $\mu = \bigoplus_{i\in I} p_i \cdot \mu_i$ there are $\bar{\nu}, \nu_i \in Distr(\mathscr{E})$, for $i \in I$, such that

$$\nu \Rightarrow \bar{\nu},\ \mu \mathscr{R} \bar{\nu},\ \bar{\nu} = \bigoplus_{i\in I} p_i \cdot \nu_i, \text{ and } \mu_i \mathscr{R} \nu_i \text{ for all } i \in I.$$

(b) A relation $\mathscr{R} \subseteq Distr(\mathscr{E}) \times Distr(\mathscr{E})$ is called a branching *probabilistic bisimulation relation iff it is weakly decomposable and for all $\mu, \nu \in Distr(\mathscr{E})$ with $\mu \mathscr{R} \nu$ and $\mu \xrightarrow{\alpha} \mu'$, there are $\bar{\nu}, \nu' \in Distr(\mathscr{E})$ such that*

$$\nu \Rightarrow \bar{\nu},\ \bar{\nu} \xrightarrow{(\alpha)} \nu',\ \mu \mathscr{R} \bar{\nu}, \text{ and } \mu' \mathscr{R} \nu'.$$

(c) Branching probabilistic bisimilarity $\underline{\leftrightarrow}_b \subseteq Distr(\mathscr{E}) \times Distr(\mathscr{E})$ is defined as the largest branching probabilistic bisimulation relation on $Distr(\mathscr{E})$.

Note that branching probabilistic bisimilarity is well-defined following the usual argument that any union of branching probabilistic bisimulation relations is again a branching probabilistic bisimulation relation. In particular, (weak) decomposability is preserved under arbitrary unions. As observed in [15], branching probabilistic bisimilarity is an equivalence relation.

Two non-deterministic processes are considered to be branching probabilistic bisimilar iff their Dirac distributions are, i.e., for $E, F \in \mathscr{E}$ we have $E \leftrightarrow_b F$ iff $\delta(E) \leftrightarrow_b \delta(F)$. Two probabilistic processes are considered to be branching probabilistic bisimilar iff their associated distributions over $\mathscr{E}$ are, i.e., for $P, Q \in \mathscr{P}$ we have $P \leftrightarrow_b Q$ iff $[\![P]\!] \leftrightarrow_b [\![Q]\!]$.

For a set $M \subseteq \mathit{Distr}(\mathscr{E})$, the convex closure $cc(M)$ is defined by

$$cc(M) = \{\textstyle\bigoplus_{i\in I} p_i\mu_i \mid \sum_{i\in I} p_i = 1,\ \mu_i \in M,\ I \text{ a finite index set}\}.$$

For a relation $\mathscr{R} \subseteq \mathit{Distr}(\mathscr{E}) \times \mathit{Distr}(\mathscr{E})$ the convex closure of $\mathscr{R}$ is defined by

$$cc(\mathscr{R}) = \{\, \langle \textstyle\bigoplus_{i\in I} p_i\mu_i, \bigoplus_{i\in I} p_i\nu_i\rangle \mid \mu_i \mathscr{R} \nu_i,\ \sum_{i\in I} p_i = 1,\ I \text{ a finite index set}\,\}.$$

The notion of weak decomposability has been adopted from [22, 24]. The underlying idea stems from [9]. Weak decomposability provides a convenient dexterity to deal with combined transitions as well as with sub-distributions. For example, regarding sub-distributions, to distinguish the probabilistic process $\frac{1}{2}\partial(a\cdot\partial(\mathbf{0})) \oplus \frac{1}{2}\partial(b\cdot\partial(\mathbf{0}))$ from $\partial(\mathbf{0})$ a branching probabilistic bisimulation relation relating $\frac{1}{2}\delta(a\cdot\partial(\mathbf{0})) \oplus \frac{1}{2}\delta(b\cdot\partial(\mathbf{0}))$ and $\delta(\mathbf{0})$ is by weak decomposability also required to relate $\delta(a\cdot\partial(\mathbf{0}))$ and $\delta(b\cdot\partial(\mathbf{0}))$ to subdistributions of a weak descendant of $\delta(\mathbf{0})$, which can only be $\delta(\mathbf{0})$ itself. Since $\delta(a\cdot\partial(\mathbf{0}))$ has an a-transition while $\delta(\mathbf{0})$ has not, and similar for a b-transition of $\delta(b\cdot\partial(\mathbf{0}))$, it follows that $\frac{1}{2}\partial(a\cdot\partial(\mathbf{0})) \oplus \frac{1}{2}\partial(b\cdot\partial(\mathbf{0}))$ and $\partial(\mathbf{0})$ are not branching probabilistic bisimilar.

By comparison, on finite processes, as used in this paper, the notion of branching probabilistic bisimilarity of Segala & Lynch [27] can be defined in our framework exactly as in (b) and (c) above, but taking a decomposable instead of a weakly decomposable relation, i.e. if $\mu \mathscr{R} \nu$ and $\mu = \bigoplus_{i\in I} p_i\mu_i$ then there are ν_i for $i \in I$ such that $\nu = \bigoplus_{i\in I} p_i\nu_i$ and $\mu_i \mathscr{R} \nu_i$ for $i \in I$. This yields a strictly finer equivalence.

Example

(a) The distributions $\delta(\mathbf{G_1}) = \delta(a\cdot(P \,{}_{\frac{1}{2}}\!\oplus Q))$ and $\delta(\mathbf{G_2}) = \delta(a\cdot(\partial(\tau\cdot(P \,{}_{\frac{1}{2}}\!\oplus Q)) \,{}_{\frac{1}{3}}\!\oplus (P \,{}_{\frac{1}{2}}\!\oplus Q)))$ both admit at the top level an a-transition only:

$$\delta(a\cdot(P \,{}_{\frac{1}{2}}\!\oplus Q)) \xrightarrow{a} \tfrac{1}{2}[\![P]\!] \oplus \tfrac{1}{2}[\![Q]\!]$$
$$\delta(a\cdot(\partial(\tau\cdot(P \,{}_{\frac{1}{2}}\!\oplus Q)) \,{}_{\frac{1}{3}}\!\oplus (P \,{}_{\frac{1}{2}}\!\oplus Q))) \xrightarrow{a} \tfrac{1}{3}\delta(\tau\cdot(P \,{}_{\frac{1}{2}}\!\oplus Q)) \oplus \tfrac{1}{3}[\![P]\!] \oplus \tfrac{1}{3}[\![Q]\!].$$

Let the relation $\mathscr{R}$ contain the pairs

$$\langle \delta(\tau\cdot(P \,{}_{\frac{1}{2}}\!\oplus Q)), \tfrac{1}{2}[\![P]\!] \oplus \tfrac{1}{2}[\![Q]\!]\rangle \quad \text{and} \quad \langle \mu, \mu\rangle \text{ for } \mu \in \mathit{Distr}(\mathscr{E}).$$

The symmetric closure $\mathscr{R}^\dagger$ of $\mathscr{R}$ is clearly a branching probabilistic bisimulation relation. We claim that therefore also its convex closure $cc(\mathscr{R}^\dagger)$ is a branching probabilistic bisimulation relation. Considering that $\langle \delta(\tau\cdot(P \,{}_{\frac{1}{2}}\!\oplus Q)), \frac{1}{2}[\![P]\!] \oplus \frac{1}{2}[\![Q]\!]\rangle$ and $\langle \frac{1}{2}[\![P]\!] \oplus \frac{1}{2}[\![Q]\!], \frac{1}{2}[\![P]\!] \oplus \frac{1}{2}[\![Q]\!]\rangle$ are in $\mathscr{R}$, we have that

$$\langle \tfrac{1}{3}\delta(\tau\cdot(P \,{}_{\frac{1}{2}}\!\oplus Q) \oplus \tfrac{2}{3}(\tfrac{1}{2}[\![P]\!] \oplus \tfrac{1}{2}[\![Q]\!]), \tfrac{1}{3}(\tfrac{1}{2}[\![P]\!] \oplus \tfrac{1}{2}[\![Q]\!]) \oplus \tfrac{2}{3}(\tfrac{1}{2}[\![P]\!] \oplus \tfrac{1}{2}[\![Q]\!]))\rangle \in cc(\mathscr{R}^\dagger).$$

Adding the pair of processes $\langle \delta(a\cdot(P \,{}_{\frac{1}{2}}\!\oplus Q)), \delta(a\cdot(\partial(\tau\cdot(P \,{}_{\frac{1}{2}}\!\oplus Q)) \,{}_{\frac{1}{3}}\!\oplus (P \,{}_{\frac{1}{2}}\!\oplus Q)))\rangle$ and closing for symmetry, then yields a branching probabilistic bisimulation relation relating $\delta(\mathbf{G_1})$ and $\delta(\mathbf{G_2})$.

(b) The a-derivatives of $\mathbf{I_1}$ and $\mathbf{I_2}$, i.e. the distributions $I_1' = \delta(b \cdot P + \tau \cdot Q)$ and $I_2' = \delta(\tau \cdot \partial(b \cdot P + \tau \cdot Q) + b \cdot P + \tau \cdot Q)$ are branching probabilistic bisimilar. A τ-transition of I_2' partially based on its left branch, can be simulated by I_1' by a partial transition:

$$\begin{array}{lll} I_2' = r \cdot [\![I_2']\!] \oplus (1-r) \cdot [\![I_2']\!] & \xrightarrow{\tau} & r \cdot \delta(b \cdot P + \tau \cdot Q) \oplus (1-r) \cdot [\![Q]\!] \\ I_1' = r \cdot [\![I_1']\!] \oplus (1-r) \cdot [\![I_1']\!] & \xrightarrow{(\tau)} & r \cdot [\![I_1']\!] \oplus (1-r) \cdot [\![Q]\!] \;=\; r \cdot \delta(b \cdot P + \tau \cdot Q) \oplus (1-r) \cdot [\![Q]\!] . \end{array}$$

A τ-transition of I_1' can be directly simulated by I_2' of course. It follows that the relation $\mathscr{R} = \{ \langle \delta(\mathbf{I_1}), \delta(\mathbf{I_2}) \rangle, \langle I_1', I_2' \rangle \}^\dagger \cup \{ \langle \mu, \mu \rangle \mid \mu \in \mathit{Distr}(\mathscr{E}) \}$, the symmetric relation containing the pairs mentioned and the diagonal of $\mathit{Distr}(\mathscr{E})$, constitutes a branching probabilistic bisimulation relation containing $\mathbf{I_1}$ and $\mathbf{I_2}$.

In the sequel we frequently need that probabilistic composition respects branching probabilistic bisimilarity of distributions, i.e. if, with respect to some index set I, we have distributions μ_i and ν_i such that $\mu_i \leftrightarroweq_b \nu_i$ for $i \in I$, then also $\mu \leftrightarroweq_b \nu$ for the distributions $\mu = \bigoplus_{i \in I} p_i \mu_i$ and $\nu = \bigoplus_{i \in I} p_i \nu_i$. The property directly follows from the following lemma, which is proven in [15].

Lemma 10. *Let distributions $\mu_1, \mu_2, \nu_1, \nu_2 \in \mathit{Distr}(\mathscr{E})$ and $0 \leqslant r \leqslant 1$ be such that $\mu_1 \leftrightarroweq_b \nu_1$ and $\mu_2 \leftrightarroweq_b \nu_2$. Then it holds that $\mu_1 \,{}_r\!\oplus\, \mu_2 \leftrightarroweq_b \nu_1 \,{}_r\!\oplus\, \nu_2$.*

We apply the above property in the proof of the next result. In the sequel any application of Lemma 10 will be done tacitly.

Lemma 11. *Let $\mu, \nu \in \mathit{Distr}(\mathscr{E})$ such that $\mu \leftrightarroweq_b \nu$ and $\mu \Rightarrow \mu'$ for some $\mu' \in \mathit{Distr}(\mathscr{E})$. Then there are $\nu' \in \mathit{Distr}(\mathscr{E})$ such that $\nu \Rightarrow \nu'$ and $\mu' \leftrightarroweq_b \nu'$.*

Proof. We check that a partial transition $\mu \xrightarrow{(\tau)} \mu'$ can be matched by ν given $\mu \leftrightarroweq_b \nu$. So, suppose $\mu = \mu_1 \,{}_r\!\oplus\, \mu_2$, $\mu_1 \xrightarrow{\tau} \mu_1'$, and $\mu' = \mu_1' \,{}_r\!\oplus\, \mu_2$. By weak decomposability of $\leftrightarroweq_b$ we can find distributions $\bar{\nu}, \nu_1, \nu_2$ such that $\nu \Rightarrow \bar{\nu} = \nu_1 \,{}_r\!\oplus\, \nu_2$ and $\mu \leftrightarroweq_b \bar{\nu}$, $\nu_1 \leftrightarroweq_b \mu_1$, $\nu_2 \leftrightarroweq_b \mu_2$. Choose distributions $\bar{\nu}_1, \bar{\nu}_1'$ such that $\nu_1 \Rightarrow \bar{\nu}_1 \xrightarrow{(\tau)} \nu_1'$ and $\bar{\nu}_1 \leftrightarroweq_b \mu_1$, $\nu_1' \leftrightarroweq_b \mu_1'$. Put $\nu' = \nu_1' \,{}_r\!\oplus\, \nu_2$. Then $\nu \Rightarrow \nu'$, using Lemma 7c, and we have by Lemma 10 that $\nu' = \nu_1' \,{}_r\!\oplus\, \nu_2 \leftrightarroweq_b \mu_1' \,{}_r\!\oplus\, \mu_2 = \mu'$ since $\nu_1' \leftrightarroweq_b \mu_1'$ and $\nu_2 \leftrightarroweq_b \mu_2$. ☐

5 Branching probabilistic bisimilarity is continuous

Fix a finite set of non-deterministic processes $\mathscr{F} \subseteq \mathscr{E}$ that is *transition closed*, in the sense that if $E \in \mathscr{F}$ and $E \xrightarrow{\alpha} \bigoplus_{i \in I} p_i \cdot F_i$ then also $F_i \in \mathscr{F}$. Consequently, if $\mu \in \mathit{Distr}(\mathscr{F})$ and $\mu \xrightarrow{(\alpha)} \mu'$ then $\mu' \in \mathit{Distr}(\mathscr{F})$. Also, if $\mu \in \mathit{Distr}(\mathscr{F})$ and $\mu \Rightarrow \bar{\mu}$ then $\bar{\mu} \in \mathit{Distr}(\mathscr{F})$. By Theorem 1 $\mathit{Distr}(\mathscr{F})$ is a sequentially compact subspace of the complete metric space $\mathit{Distr}(\mathscr{E})$, meaning that every sequence $(\mu_i)_{i=0}^\infty$ in $\mathit{Distr}(\mathscr{F})$ has a subsequence $(\mu_{i_k})_{k=0}^\infty$ such that $\lim_{k \to \infty} \mu_{i_k} = \mu$ for some distribution $\mu \in \mathit{Distr}(\mathscr{F})$. In particular, if $\lim_{i \to \infty} \mu_i = \mu$ and $\mu_i \in \mathit{Distr}(\mathscr{F})$, then also $\mu \in \mathit{Distr}(\mathscr{F})$, i.e. $\mathit{Distr}(\mathscr{F})$ is a closed subset of $\mathit{Distr}(\mathscr{E})$. Due to the finitary nature of our process algebra, each distribution $\mu \in \mathit{Distr}(\mathscr{E})$ occurs in $\mathit{Distr}(\mathscr{F})$ for some such $\mathscr{F}$, based on $\mathit{spt}(\mu)$.

In the following three lemmas we establish a number of continuity results. Assume $\lim_{i \to \infty} \nu_i = \nu$. Then Lemma 12 states that, for a Dirac distribution $\delta(E)$, if $\delta(E) \xrightarrow{\alpha} \nu_i$ for $i \in \mathbb{N}$ then also $\delta(E) \xrightarrow{\alpha} \nu$. Lemma 13 extends this and shows that, for a general distribution μ, if $\mu \xrightarrow{\alpha} \nu_i$ for $i \in \mathbb{N}$ then $\mu \xrightarrow{\alpha} \nu$. Finally, Lemma 14 establishes the limit case: if $\lim_{i \to \infty} \mu_i = \mu$ and $\mu_i \xrightarrow{\alpha} \nu_i$ for $i \in \mathbb{N}$ then $\mu \xrightarrow{\alpha} \nu$.

Lemma 12. *Let $E \in \mathscr{F}$ be a non-deterministic process, $\alpha \in \mathscr{A}$ an action, $(\nu_i)_{i=0}^\infty \in \mathit{Distr}(\mathscr{F})^\infty$ an infinite sequence in $\mathit{Distr}(\mathscr{F})$, and $\nu \in \mathit{Distr}(\mathscr{F})$ a distribution satisfying $\lim_{i \to \infty} \nu_i = \nu$. If, for all $i \in \mathbb{N}$, $\delta(E) \xrightarrow{(\alpha)} \nu_i$ then it holds that $\delta(E) \xrightarrow{(\alpha)} \nu$.*

Proof. For $E \in \mathscr{F}$ and $\alpha \in \mathscr{A}$, define $E\restriction\alpha = cc(\{\, \mu \mid E \xrightarrow{\alpha} \mu\})$, pronounced E 'after' α, to be the convex closure in $Distr(\mathscr{E})$ of all distributions that can be reached from E by an α-transition. Then $\delta(E) \xrightarrow{\alpha} \nu$ iff $\nu \in E\restriction\alpha$. Recall that transitions for non-deterministic processes are not probabilistically combined. See Definition 5. Since $E\restriction\alpha \subseteq Distr(\mathscr{F})$ is the convex closure of a finite set of distributions, it is certainly closed in the space $Distr(\mathscr{F})$. Since it holds that $\delta(E) \xrightarrow{\alpha} \nu_i$ for all $i \in \mathbb{N}$, one has $\nu_i \in E\restriction\alpha$ for $i \in \mathbb{N}$. Hence, $\lim_{i\to\infty} \nu_i = \nu$ implies that $\nu \in E\restriction\alpha$, i.e. $\delta(E) \xrightarrow{\alpha} \nu$.

For $E \in \mathscr{F}$, define $E\restriction(\tau) := cc(\{\mu \mid E \xrightarrow{\tau} \mu\} \cup \{E\})$. Then $\delta(E) \xrightarrow{(\tau)} \nu$ iff $\nu \in E\restriction(\tau)$. The set $E\restriction(\tau) \subseteq Distr(\mathscr{F})$ is closed, and thus $\nu_i \in E\restriction(\tau)$ implies $\nu \in E\restriction(\tau)$, which means $\delta(E) \xrightarrow{(\tau)} \nu$. □

The above result for Dirac distributions holds for general distributions as well.

Lemma 13. *Let $\mu, \nu \in Distr(\mathscr{F})$, $\alpha \in \mathscr{A}$, $(\nu_i)_{i=0}^{\infty} \in Distr(\mathscr{F})^{\infty}$, and assume $\lim_{i\to\infty} \nu_i = \nu$. If it holds that $\mu \xrightarrow{(\alpha)} \nu_i$ for all $i \in \mathbb{N}$, then also $\mu \xrightarrow{(\alpha)} \nu$.*

Proof. Suppose $\mu \xrightarrow{(\alpha)} \nu_i$ for all $i \in I$. Let $\mu = \bigoplus_{j=1}^{k} p_j \cdot E_j$. By Lemma 8, for all $i \in \mathbb{N}$ and $1 \leqslant j \leqslant k$ there are ν_{ij} such that $\delta(E_j) \xrightarrow{(\alpha)} \nu_{ij}$ and $\nu_i = \bigoplus_{j=1}^{k} p_j \cdot \nu_{ij}$. The countable sequence $(\nu_{i1}, \nu_{i2}, \ldots, \nu_{ik})_{i=0}^{\infty}$ of k-dimensional vectors of probability distributions need not have a limit. However, by the sequential compactness of $Distr(\mathscr{F})$ this sequence has an infinite subsequence in which the first components ν_{i_1} converge to a limit η_1. That sequence in turn has an infinite subsequence in which also the second components ν_{i_2} converge to a limit η_2. Going on this way, one finds a subsequence $(\nu_{i_h1}, \nu_{i_h2}, \ldots, \nu_{i_hk})_{h=0}^{\infty}$ of $(\nu_{i1}, \nu_{i2}, \ldots, \nu_{ik})_{i=0}^{\infty}$ for $i_0 < i_1 < \ldots$ that has a limit, say $\lim_{h\to\infty}(\nu_{i_h1}, \nu_{i_h2}, \ldots, \nu_{i_hk}) = (\eta_1, \eta_2, \ldots, \eta_k)$. Using that $\lim_{h\to\infty} \nu_{i_h} = \nu$, one obtains $\nu = \bigoplus_{j=1}^{k} p_j \cdot \eta_j$. For each $j = 1, \ldots, k$, by Lemma 12, since $\delta(E_j) \xrightarrow{(\alpha)} \nu_{ij}$ for all $i \in I$ and $\lim_{h\to\infty} \nu_{i_hj} = \eta_j$, we conclude that $\delta(E_j) \xrightarrow{(\alpha)} \eta_j$. Thus, by Lemma 7, $\mu = \bigoplus_{j=1}^{k} p_j \cdot E_j \xrightarrow{(\alpha)} \bigoplus_{j=1}^{k} p_j \cdot \eta_j = \nu$. □

Next, we consider a partial transition over a convergent sequence of distributions.

Lemma 14. *Let $(\mu_i)_{i=0}^{\infty}, (\nu_i)_{i=0}^{\infty} \in Distr(\mathscr{F})^{\infty}$ such that $\lim_{i\to\infty} \mu_i = \mu$ and $\lim_{i\to\infty} \nu_i = \nu$. If it holds that $\mu_i \xrightarrow{(\alpha)} \nu_i$ for all $i \in \mathbb{N}$, then also $\mu \xrightarrow{(\alpha)} \nu$.*

Proof. Since $\lim_{i\to\infty} \mu_i = \mu$, we can write $\mu_i = (1-r_i)\mu \oplus r_i\mu_i''$, for suitable $\mu_i'' \in Distr(\mathscr{F})$ and $r_i \geqslant 0$ such that $\lim_{i\to\infty} r_i = 0$, as guaranteed by Lemma 2. Because $\mu_i \xrightarrow{(\alpha)} \nu_i$, by Lemma 8 there are distributions $\nu_i', \nu_i'' \in Distr(\mathscr{F})$ for $i \in \mathbb{N}$ such that $\nu_i = (1-r_i)\nu_i' \oplus r_i\nu_i''$, $\mu \xrightarrow{(\alpha)} \nu_i'$, and $\mu_i'' \xrightarrow{(\alpha)} \nu_i''$. We have $\lim_{i\to\infty} \nu_i' = \nu$ as well, since $\lim_{i\to\infty} r_i = 0$. Thus, $\lim_{i\to\infty} \nu_i' = \nu$ and $\mu \xrightarrow{(\alpha)} \nu_i'$ for $i \in \mathbb{N}$. Therefore, it follows by Lemma 13 that $\mu \xrightarrow{(\alpha)} \nu$. □

For $\mu, \nu \in Distr(\mathscr{F})$, we write $\mu \Rightarrow_n \nu$ if there are $\eta_0, \eta_1, \ldots, \eta_n \in Distr(\mathscr{F})$ such that $\mu = \eta_0 \xrightarrow{(\tau)} \eta_1 \xrightarrow{(\tau)} \ldots \xrightarrow{(\tau)} \eta_n = \nu$. Clearly, it holds that $\mu \Rightarrow_n \nu$ for some $n \in \mathbb{N}$ in case $\mu \Rightarrow \nu$, because $\Rightarrow$ is the transitive closure of $\xrightarrow{(\tau)}$.

We have the following pendant of Lemma 14 for $\Rightarrow_n$.

Lemma 15. *Let $(\mu_i)_{i=0}^{\infty}, (\nu_i)_{i=0}^{\infty} \in Distr(\mathscr{F})^{\infty}$, $\lim_{i\to\infty} \mu_i = \mu$ and $\lim_{i\to\infty} \nu_i = \nu$. If $\mu_i \Rightarrow_n \nu_i$ for all $i \in \mathbb{N}$ then $\mu \Rightarrow_n \nu$.*

Proof. By induction on n. Basis, $n = 0$: Trivial. Induction step, $n+1$: Given $(\mu_i)_{i=0}^{\infty}, (\nu_i)_{i=0}^{\infty} \in Distr(\mathscr{F})^{\infty}$, $\mu = \lim_{i\to\infty} \mu_i$, and $\nu = \lim_{i\to\infty} \nu_i$, suppose $\mu_i \Rightarrow_{n+1} \nu_i$ for all $i \in \mathbb{N}$. Let $(\eta_i)_{i=0}^{\infty} \in Distr(\mathscr{F})^{\infty}$ be such that $\mu_i \xrightarrow{(\tau)} \eta_i \Rightarrow_n \nu_i$ for all $i \in \mathbb{N}$. Since $Distr(\mathscr{F})$ is sequentially compact, the sequence $(\eta_i)_{i=0}^{\infty}$ has a convergent subsequence $(\eta_{i_k})_{k=0}^{\infty}$; put $\eta = \lim_{k\to\infty} \eta_{i_k}$. Because $\mu_{i_k} \xrightarrow{(\tau)} \eta_{i_k}$ for all $k \in \mathbb{N}$, one has $\mu \xrightarrow{(\tau)} \eta$ by Lemma 14. Since $\eta_{i_k} \Rightarrow_n \nu_{i_k}$ for $k \in \mathbb{N}$, the induction hypothesis yields $\eta \Rightarrow_n \nu$. It follows that $\mu \Rightarrow_{n+1} \nu$. □

We adapt Lemma 15 to obtain a continuity result for weak transitions $\Rightarrow$.

Lemma 16. *Let* $(\mu_i)_{i=0}^{\infty}, (\nu_i)_{i=0}^{\infty} \in \mathit{Distr}(\mathscr{F})^{\infty}$, $\lim_{i\to\infty} \mu_i = \mu$ *and* $\lim_{i\to\infty} \nu_i = \nu$. *If* $\mu_i \Rightarrow \nu_i$ *for all* $i \in \mathbb{N}$, *then* $\mu \Rightarrow \nu$.

Proof. Since $\mathscr{F}$ contains only finitely many non-deterministic processes, which can do finitely many τ-transitions only, a global upperbound N exists such that if $\mu \Rightarrow \nu$ then $\mu \Rightarrow_k \nu$ for some $k \leqslant N$.

Moreover, as each sequence $\mu = \eta_0 \xrightarrow{(\tau)} \eta_1 \xrightarrow{(\tau)} \ldots \xrightarrow{(\tau)} \eta_k = \nu$ with $k < N$ can be extended to a sequence $\mu = \eta_0 \xrightarrow{(\tau)} \eta_1 \xrightarrow{(\tau)} \ldots \xrightarrow{(\tau)} \eta_N = \nu$, namely by taking $\eta_i = \nu$ for all $k < i \leqslant N$, on $\mathscr{F}$ the relations $\Rightarrow$ and $\Rightarrow_N$ coincide. Consequently, Lemma 16 follows from Lemma 15. $\square$

The following theorem says that equivalence classes of branching probabilistic bisimilarity in $\mathit{Distr}(\mathscr{F})$ are closed sets of distributions.

Theorem 17. *Let* $\hat{\mu}, \hat{\nu} \in \mathit{Distr}(\mathscr{F})$ *and* $(\nu_i)_{i=0}^{\infty} \in \mathit{Distr}(\mathscr{F})^{\infty}$ *such that* $\hat{\mu} \leftrightarrow_b \nu_i$ *for all* $i \in \mathbb{N}$ *and* $\hat{\nu} = \lim_{i\to\infty} \nu_i$. *Then it holds that* $\hat{\mu} \leftrightarrow_b \hat{\nu}$.

Proof. Define the relation $\mathscr{R}$ on $\mathit{Distr}(\mathscr{F})$ by

$$\begin{aligned} \mu \,\mathscr{R}\, \nu \iff\ & \exists (\mu_i)_{i=0}^{\infty}, (\nu_i)_{i=0}^{\infty} \in \mathit{Distr}(\mathscr{F})^{\infty} : \\ & \lim_{i\to\infty} \mu_i = \mu \wedge \lim_{i\to\infty} \nu_i = \nu \wedge \forall i \in \mathbb{N} \colon \mu_i \leftrightarrow_b \nu_i \end{aligned}$$

As $\hat{\mu} \,\mathscr{R}\, \hat{\nu}$ (taking $\mu_i := \hat{\mu}$ for all $i \in I$), it suffices to show that $\mathscr{R}$ is a branching probabilistic bisimulation.

Suppose $\mu \,\mathscr{R}\, \nu$. Let $(\mu_i)_{i=0}^{\infty}, (\nu_i)_{i=0}^{\infty} \in \mathit{Distr}(\mathscr{F})^{\infty}$ be such that $\lim_{i\to\infty} \mu_i = \mu$, $\lim_{i\to\infty} \nu_i = \nu$, and $\mu_i \leftrightarrow_b \nu_i$ for all $i \in \mathbb{N}$. Since $\lim_{i\to\infty} \mu_i = \mu$, there exist $(\mu_i')_{i=0}^{\infty} \in \mathit{Distr}(\mathscr{F})^{\infty}$ and $(r_i)_{i=0}^{\infty} \in \mathbb{R}^{\infty}$ such that $\mu_i = (1-r_i)\mu \oplus r_i \mu_i'$ for all $i \in \mathbb{N}$ and $\lim_{i\to\infty} r_i = 0$.

(i) Towards weak decomposability of $\mathscr{R}$ for μ vs. ν, suppose $\mu = \bigoplus_{j\in J} q_j \cdot \bar{\mu}_j$. So, for all $i \in \mathbb{N}$, we have that $\mu_i = (1-r_i)\big(\bigoplus_{j\in J} q_j \cdot \bar{\mu}_j\big) \oplus r_i \mu_i'$. By weak decomposability of $\leftrightarrow_b$, there exist $\bar{\bar{\nu}}_i$, ν_i' and ν_{ij} for $i \in \mathbb{N}$ and $j \in J$ such that $\nu_i \Rightarrow \bar{\bar{\nu}}_i$, $\mu_i \leftrightarrow_b \bar{\bar{\nu}}_i$, $\bar{\bar{\nu}}_i = (1-r_i)\big(\bigoplus_{j\in J} q_j \cdot \nu_{ij}\big) \oplus r_i \nu_i'$, $\mu_i' \leftrightarrow_b \nu_i'$, and $\bar{\mu}_j \leftrightarrow_b \nu_{ij}$ for $j \in J$.

The sequences $(\nu_{ij})_{i=0}^{\infty}$ for $j \in J$ may not converge. However, by sequential compactness of $\mathit{Distr}(\mathscr{F})$ (and successive sifting out for each $j \in J$) an index sequence $(i_k)_{k=0}^{\infty}$ exists such that the sequences $(\nu_{i_k j})_{k=0}^{\infty}$ converge, say $\lim_{k\to\infty} \nu_{i_k j} = \bar{\nu}_j$ for $j \in J$. Put $\bar{\nu} = \bigoplus_{j\in J} q_j \cdot \bar{\nu}_j$. Then it holds that

$$\lim_{k\to\infty} \bar{\bar{\nu}}_{i_k} = \lim_{k\to\infty} (1-r_{i_k})\big(\textstyle\bigoplus_{j\in J} q_j \cdot \nu_{i_k j}\big) \oplus r_{i_k} \nu_{i_k}' = \lim_{k\to\infty} \textstyle\bigoplus_{j\in J} q_j \cdot \nu_{i_k j} = \textstyle\bigoplus_{j\in J} q_j \cdot \bar{\nu}_j = \bar{\nu}$$

as $\lim_{k\to\infty} r_{i_k} = 0$ and probabilistic composition is continuous. Since $\nu_{i_k} \Rightarrow \bar{\bar{\nu}}_{i_k}$ for all $k \in \mathbb{N}$, one has $\lim_{k\to\infty} \nu_{i_k} \Rightarrow \lim_{k\to\infty} \bar{\bar{\nu}}_{i_k}$, i.e. $\nu \Rightarrow \bar{\nu}$, by Lemma 16. Also, $\mu_{i_k} \leftrightarrow_b \bar{\bar{\nu}}_{i_k}$ for all $k \in \mathbb{N}$. Therefore, by definition of $\mathscr{R}$, we obtain $\mu \,\mathscr{R}\, \bar{\nu}$. Since $\bar{\mu}_j \leftrightarrow_b \nu_{i_k j}$ for all $k \in \mathbb{N}$ and $j \in J$, it follows that $\bar{\mu}_j \,\mathscr{R}\, \bar{\nu}_j$ for $j \in J$. Thus, $\nu \Rightarrow \bar{\nu} = \bigoplus_{j\in J} q_j \cdot \bar{\nu}_j$, $\mu \,\mathscr{R}\, \bar{\nu}$, and $\bar{\mu}_j \,\mathscr{R}\, \bar{\nu}_j$ for all $j \in J$, as was to be shown. Hence the relation $\mathscr{R}$ is weakly decomposable.

(ii) For the transfer property, suppose $\mu \xrightarrow{\alpha} \mu'$ for some $\alpha \in \mathscr{A}$. Since, for each $i \in \mathbb{N}$, $\mu_i \leftrightarrow_b \nu_i$ and $\mu_i = (1-r_i)\mu \oplus r_i \mu_i'$, it follows from weak decomposability of $\leftrightarrow_b$ that distributions $\bar{\nu}_i$, ν_i' and ν_i'' exist such that $\nu_i \Rightarrow \bar{\nu}_i$, $\mu_i \leftrightarrow_b \bar{\nu}_i$, $\bar{\nu}_i = (1-r_i)\nu_i' \oplus r_i \nu_i''$ and $\mu \leftrightarrow_b \nu_i'$. By the transfer property for $\leftrightarrow_b$, for each $i \in \mathbb{N}$ exist $\bar{\eta}_i, \eta_i' \in \mathit{Distr}(\mathscr{E})$ such that

$$\nu_i' \Rightarrow \bar{\eta}_i,\ \bar{\eta}_i \xrightarrow{(\alpha)} \eta_i',\ \mu \leftrightarrow_b \bar{\eta}_i, \text{ and } \mu' \leftrightarrow_b \eta_i'.$$

We have $\bar{\nu}_i' \in \mathit{Distr}(\mathscr{F})$ for $i \in \mathbb{N}$. Also, $\bar{\eta}_i, \eta_i' \in \mathit{Distr}(\mathscr{F})$ for $i \in \mathbb{N}$, since $\mathscr{F}$ is assumed to be transition closed. Therefore, by sequential compactness of $\mathit{Distr}(\mathscr{F})$, the sequences $(\bar{\nu}_i')_{i=0}^{\infty}$, $(\bar{\eta}_i)_{i=0}^{\infty}$,

$(\bar{\eta}_i')_{i=0}^\infty$ have converging subsequences $(\bar{\nu}_{i_k}')_{k=0}^\infty$, $(\bar{\eta}_{i_k})_{k=0}^\infty$, and $(\bar{\eta}_{i_k}')_{k=0}^\infty$, respectively. Put $\bar{\nu} = \lim_{k\to\infty} \nu_{i_k}'$, $\bar{\eta} = \lim_{k\to\infty} \bar{\eta}_{i_k}$, and $\eta' = \lim_{k\to\infty} \eta_{i_k}'$. As $\lim_{k\to\infty} r_{i_k} = 0$, one has $\lim_{k\to\infty} \bar{\nu}_{i_k} = \lim_{k\to\infty} \nu_{i_k}' = \bar{\nu}$.

Since $\nu_{i_k} \Rightarrow \bar{\nu}_{i_k}$ for $k \in \mathbb{N}$, we obtain $\lim_{k\to\infty} \nu_{i_k} \Rightarrow \lim_{k\to\infty} \bar{\nu}_{i_k}$ by Lemma 16, thus $\nu \Rightarrow \bar{\nu}$. Likewise, as $\nu_{i_k}' \Rightarrow \bar{\eta}_{i_k}$ for all $k \in \mathbb{N}$, one has $\bar{\nu} \Rightarrow \bar{\eta}$, and therefore $\nu \Rightarrow \bar{\eta}$. Furthermore, because $\bar{\eta}_{i_k} \xrightarrow{(\alpha)} \eta_{i_k}'$ for $k \in \mathbb{N}$, it follows that $\bar{\eta} \xrightarrow{(\alpha)} \eta'$, now by Lemma 14. From $\mu \leftrightarroweq_b \bar{\eta}_{i_k}$ for all $k \in \mathbb{N}$, we obtain $\mu \mathscr{R} \bar{\eta}$ by definition of $\mathscr{R}$. Finally, $\mu' \leftrightarroweq_b \eta_{i_k}'$ for all $k \in \mathbb{N}$ yields $\mu' \mathscr{R} \eta'$. Thus $\nu \Rightarrow \bar{\eta} \xrightarrow{(\alpha)} \eta'$, $\mu \mathscr{R} \bar{\eta}$, and $\mu' \mathscr{R} \bar{\eta}'$, which was to be shown. □

The following corollary of Theorem 17 will be used in the next section.

Corollary 18. *For each $\mu \in Distr(\mathscr{E})$, the set $T_\mu = \{ \nu \in Distr(\mathscr{E}) \mid \nu \leftrightarroweq_b \mu \wedge \mu \Rightarrow \nu \}$ is a sequentially compact set.*

Proof. For $\mu = \bigoplus_{i\in I} p_i \cdot E_i$, the set of processes $\mathscr{F} = \{ E \in \mathscr{E} \mid E$ occurs in E_i for some $i \in I \}$ is finite and closed under transitions. Clearly, $\mu \in Distr(\mathscr{F})$. Moreover, $Distr(\mathscr{F})$ is a sequentially compact subset of $Distr(\mathscr{E})$. Taking $\mu_i = \mu$ for all $i \in \mathbb{N}$ in Lemma 16 yields that $\{ \nu \mid \mu \Rightarrow \nu \}$ is a closed subset of $Distr(\mathscr{F})$. Similarly, the set $\{ \nu \mid \nu \leftrightarroweq_b \mu \}$ is a closed subset of $Distr(\mathscr{F})$ by Theorem 17. The statement then follows since the intersection of two closed subsets of $Distr(\mathscr{F})$ is itself closed, and hence sequentially compact. □

6 Cancellativity for branching probabilistic bisimilarity

With the results of Section 5 in place, we turn to stable processes and cancellativity. In the introduction we argued that in general it doesn't need to be the case that two branching probabilistic bisimilar distributions assign the same weight to equivalence classes. Here we show that this property does hold when restricting to stable distributions. We continue to prove the announced unfolding result, that for every distribution μ there exists a stable distribution σ such that $\mu \Rightarrow \sigma$ and $\mu \leftrightarroweq_b \sigma$. That result will be pivotal in the proof of the cancellation theorem, Theorem 22.

Definition 19. *A distribution $\mu \in Distr(\mathscr{E})$ is called* stable *if, for all $\bar{\mu} \in Distr(\mathscr{E})$, $\mu \Rightarrow \bar{\mu}$ and $\mu \leftrightarroweq_b \bar{\mu}$ imply that $\bar{\mu} = \mu$.*

Thus, a distribution μ is called stable if it cannot perform internal activity without leaving its branching bisimulation equivalence class. By definition of $\xrightarrow{(\tau)}$ it is immediate that if $\bigoplus_{i\in I} p_i \cdot \mu_i$ is a stable distribution with $p_i > 0$ for $i \in I$, then also each probabilistic component μ_i is stable. Also, because two stable distributions μ and ν don't have any non-trivial partial τ-transitions, weak decomposability between them amounts to decomposability, i.e. if $\mu \leftrightarroweq_b \nu$ and $\mu = \bigoplus_{i\in I} p_i \mu_i$ then distributions ν_i for $i \in I$ exist such that $\nu = \bigoplus_{i\in I} p_i \nu_i$ and $\mu_i \leftrightarroweq_b \nu_i$ for $i \in I$.

The next result states that, contrary to distributions in general, two stable distributions are branching bisimilar precisely when they assign the same probability on all branching bisimilarity classes of $\mathscr{E}$.

Lemma 20. *Let $\mu, \nu \in Distr(\mathscr{E})$ be two stable distributions. Then it holds that $\mu \leftrightarroweq_b \nu$ iff $\mu[C] = \nu[C]$ for each equivalence class C of branching probabilistic bisimilarity in $\mathscr{E}$.*

Proof. Suppose $\mu = \bigoplus_{i\in I} p_i \cdot E_i$, $\nu = \bigoplus_{j\in J} q_j \cdot F_j$, and $\mu \leftrightarroweq_b \nu$. By weak decomposability, $\nu \Rightarrow \bar{\nu} = \bigoplus_{i\in I} p_i \cdot \nu_i$ for suitable $\nu_i \in Distr(\mathscr{E})$ for $i \in I$ with $\nu_i \leftrightarroweq_b \delta(E_i)$ and $\bar{\nu} \leftrightarroweq_b \mu$. Hence, $\bar{\nu} \leftrightarroweq_b \mu \leftrightarroweq_b \nu$. Thus, by stability of ν, we have $\bar{\nu} = \nu$. Say, $\nu_i = \bigoplus_{j\in J} q_{ij} \cdot F_j$ with $q_{ij} \geqslant 0$, for $i \in I$, $j \in J$. Since $\nu_i \leftrightarroweq_b \delta(E_i)$, we have by weak decomposability, $\delta(E_i) \Rightarrow \bigoplus_{j\in J} q_{ij} \cdot \mu_{ij}'$ such that $\delta(E_i) \leftrightarroweq_b \bigoplus_{j\in J} q_{ij} \cdot \mu_{ij}'$

and $\mu'_{ij} \leftrightarroweq_b \delta(F_j)$ for suitable $\mu'_{ij} \in Distr(\mathscr{E})$. Since μ is stable, so is $\delta(E_i)$. Hence $\delta(E_i) = \bigoplus_{j\in J} q_{ij}\cdot\mu'_{ij}$, $\mu'_{ij} = \delta(E_i)$, and $E_i \leftrightarroweq_b F_j$ if $q_{ij} > 0$. Put $p_{ij} = p_i q_{ij}$, $E_{ij} = E_i$ if $q_{ij} > 0$, and $E_{ij} = \mathbf{0}$ otherwise, $F_{ij} = F_j$ if $q_{ij} > 0$, and $F_{ij} = \mathbf{0}$ otherwise, for $i \in I$, $j \in J$. Then it holds that

$$\begin{aligned} \mu &= \textstyle\bigoplus_{i\in I} p_i \cdot E_i = \bigoplus_{i\in I} p_i \cdot \big(\bigoplus_{j\in J} q_{ij}\cdot E_i\big) = \bigoplus_{i\in I}\bigoplus_{j\in J} p_i q_{ij}\cdot E_i = \bigoplus_{i\in I}\bigoplus_{j\in J} p_{ij}\cdot E_{ij} \\ \nu &= \textstyle\bigoplus_{i\in I} p_i \cdot \nu_i = \bigoplus_{i\in I} p_i \cdot \big(\bigoplus_{j\in J} q_{ij}\cdot F_j\big) = \bigoplus_{i\in I}\bigoplus_{j\in J} p_i q_{ij}\cdot F_j = \bigoplus_{i\in I}\bigoplus_{j\in J} p_{ij}\cdot F_{ij}. \end{aligned}$$

Now, for any equivalence class C of $\mathscr{E}$ modulo $\leftrightarroweq_b$, it holds that $E_{ij} \in C \Leftrightarrow F_{ij} \in C$ for all indices $i \in I$, $j \in J$. So, $\mu[C] = \sum_{i\in I, j\in J:\, E_{ij}\in C} p_{ij} = \sum_{i\in I, j\in J:\, F_{ij}\in C} p_{ij} = \nu[C]$.

For the reverse direction, suppose $\mu = \bigoplus_{i\in I} p_i \cdot E_i$, $\nu = \bigoplus_{j\in J} q_j \cdot F_j$, with $p_i, q_j > 0$, and $\mu[C] = \nu[C]$ for each equivalence class $C \in \mathscr{E}/\leftrightarroweq$.

For $i \in I$ and $j \in J$, let C_i and D_j be the equivalence class in $\mathscr{E}$ of E_i and F_j modulo $\leftrightarroweq_b$. Define $r_{ij} = \delta_{ij} p_i q_j / \mu[C_i]$, for $i \in I$, $j \in J$, where $\delta_{ij} = 1$ if $E_i \leftrightarroweq_b F_j$ and $\delta_{ij} = 0$ otherwise. Then it holds that

$$\textstyle\sum_{j\in J} r_{ij} = \sum_{j\in J} \dfrac{\delta_{ij} p_i q_j}{\mu[C_i]} = \dfrac{p_i}{\mu[C_i]} \sum_{j\in J} \delta_{ij} q_j = \dfrac{p_i \nu[C_i]}{\mu[C_i]} = p_i.$$

Since $\delta_{ij} p_i q_j / \mu[C_i] = \delta_{ij} p_i q_j / \nu[D_j]$ for $i \in I$, $j \in J$, we also have $\sum_{i\in I} r_{ij} = q_j$. Therefore, we can write $\mu = \bigoplus_{i\in I}\bigoplus_{j\in J} r_{ij}\cdot E_{ij}$ and $\nu = \bigoplus_{i\in I}\bigoplus_{j\in J} r_{ij}\cdot F_{ij}$ for suitable E_{ij} and F_{ij} such that $E_{ij} \leftrightarroweq_b F_{ij}$. Calling Lemma 10 it follows that $\mu \leftrightarroweq_b \nu$. □

Next, in Lemma 21, we are about to prove a crucial property for our proof of cancellativity, the proof of Theorem 22 below. Generally, a distribution may allow inert partial transitions. However, the distribution can be unfolded to reach via inert partial transitions a stable distribution, which doesn't have these by definition. To obtain the result we will rely on the topological property of sequential compactness of the set $T_\mu = \{\, \mu' \mid \mu' \leftrightarroweq_b \mu \wedge \mu \Rightarrow \mu' \,\}$ introduced in the previous section.

Lemma 21. *For all $\mu \in Distr(\mathscr{E})$ there is a stable distribution $\sigma \in Distr(\mathscr{E})$ such that $\mu \Rightarrow \sigma$.*

Proof. Define the *weight* of a distribution by $wgt(\mu) = \sum_{E\in\mathscr{E}} \mu(E)\cdot c(E)$, i.e., the weighted average of the complexities of the states in its support. In view of these definitions, $E \xrightarrow{\alpha} \mu$ implies $wgt(\mu) < wgt(\delta(E))$ and $\mu \xrightarrow{\alpha} \mu'$ implies $wgt(\mu') < wgt(\mu)$. In addition, $\mu \Rightarrow \mu'$ implies $wgt(\mu') \leqslant wgt(\mu)$.

For a distribution $\mu \in Distr(\mathscr{E})$, the set T_μ is given by $T_\mu = \{\, \mu' \mid \mu' \leftrightarroweq_b \mu \wedge \mu \Rightarrow \mu' \,\}$. Consider the value $\inf\{\, wgt(\mu') \mid \mu' \in T_\mu \,\}$. By Corollary 18, T_μ is a sequentially compact set. Since the infimum over a sequentially compact set will be reached, there exists a distribution σ such that $\mu \Rightarrow \sigma$, $\sigma \leftrightarroweq_b \mu$, and $wgt(\sigma) = \inf\{\, wgt(\mu') \mid \mu' \in T_\mu \,\}$. By definition of T_μ, the distribution σ must be stable. □

We have arrived at the main result of the paper, slightly more general formulated compared to the description in the introduction. The message remains the same: if two distributions are branching probabilistic bisimilar and have components that are branching probabilistic bisimilar, then the components that remain after cancelling the earlier components are also branching probabilistic bisimilar. As we see, the previous lemma is essential in the proof as given.

Theorem 22 (Cancellativity)**.** *Let $\mu, \mu', \nu, \nu' \in Distr(\mathscr{E})$ and $0 < r \leqslant 1$ be such that $\mu \,{}_r\!\oplus\, \nu \leftrightarroweq_b \mu' \,{}_r\!\oplus\, \nu'$ and $\nu \leftrightarroweq_b \nu'$. Then it holds that $\mu \leftrightarroweq_b \mu'$.*

Proof. Choose μ, μ', ν, ν', and r according to the premise of the theorem. By Lemma 21, a stable distribution σ exists such that $\mu \,{}_r\!\oplus\, \nu \Rightarrow \sigma$ and $\sigma \leftrightarroweq_b \mu \,{}_r\!\oplus\, \nu$. By weak decomposability, we can find distributions $\bar{\mu}$ and $\bar{\nu}$ such that $\sigma \Rightarrow \bar{\mu} \,{}_r\!\oplus\, \bar{\nu}$, $\bar{\mu} \leftrightarroweq_b \mu$, and $\bar{\nu} \leftrightarroweq_b \nu$. By stability of σ we have $\sigma = \bar{\mu} \,{}_r\!\oplus\, \bar{\nu}$.

Thus $\bar{\mu}\;{}_r\!\oplus\;\bar{\nu}$ is stable. Symmetrically, there are distributions $\bar{\mu}'$ and $\bar{\nu}'$ such that $\bar{\mu}' \leftrightarroweq_b \mu'$, $\bar{\nu}' \leftrightarroweq_b \nu'$ and such that $\bar{\mu}'\;{}_r\!\oplus\;\bar{\nu}'$ is stable. Note, $\bar{\mu}\;{}_r\!\oplus\;\bar{\nu} \leftrightarroweq_b \mu\;{}_r\!\oplus\;\nu \leftrightarroweq_b \mu'\;{}_r\!\oplus\;\nu' \leftrightarroweq_b \bar{\mu}'\;{}_r\!\oplus\;\bar{\nu}'$.

Let $C \subseteq \mathcal{E}$ be an equivalence class of $\mathcal{E}/\leftrightarroweq_b$. The distributions $\bar{\mu}\;{}_r\!\oplus\;\bar{\nu}$ and $\bar{\mu}'\;{}_r\!\oplus\;\bar{\nu}'$ are stable and $\bar{\mu}\;{}_r\!\oplus\;\bar{\nu} \leftrightarroweq_b \bar{\mu}'\;{}_r\!\oplus\;\bar{\nu}'$. From Lemma 20 we obtain that $(\bar{\mu}\;{}_r\!\oplus\;\bar{\nu})[C] = (\bar{\mu}'\;{}_r\!\oplus\;\bar{\nu}')[C]$. Since ν and $\bar{\nu}$ are stable and $\bar{\nu} \leftrightarroweq_b \bar{\nu}'$, we have $\bar{\nu}[C] = \bar{\nu}'[C]$ for the same reason. Because $(\bar{\mu}\;{}_r\!\oplus\;\bar{\nu})[C] = r \cdot \bar{\mu}[C] + (1-r) \cdot \bar{\nu}[C]$ and $(\bar{\mu}'\;{}_r\!\oplus\;\bar{\nu}')[C] = r \cdot \bar{\mu}'[C] + (1-r) \cdot \bar{\nu}'[C]$, we calculate

$$r \cdot \bar{\mu}[C] = (\bar{\mu}\;{}_r\!\oplus\;\bar{\nu})[C] - (1-r) \cdot \bar{\nu}[C] = (\bar{\mu}'\;{}_r\!\oplus\;\bar{\nu}')[C] - (1-r) \cdot \bar{\nu}'[C] = r \cdot \bar{\mu}'[C].$$

Since $r \neq 0$, it follows $\bar{\mu}[C] = \bar{\mu}'[C]$. Since $\bar{\mu}$ and $\bar{\mu}'$ are stable it follows by Lemma 20 that $\bar{\mu} \leftrightarroweq_b \bar{\mu}'$. Consequently, $\mu \leftrightarroweq_b \bar{\mu} \leftrightarroweq_b \bar{\mu}' \leftrightarroweq_b \mu'$. In particular $\mu \leftrightarroweq_b \mu'$, as was to be shown. □

7 Concluding remarks

We have shown a cancellation law for distributions with respect to branching probabilistic bisimilarity. The result rests on the notion of a stable distribution. Stable distributions enjoy two properties that have been essential to our set-up. (i) Every distribution has a weak unfolding towards a stable distribution that is branching probabilistic bisimilar. (ii) Branching probabilistic bisimilarity for stable distributions is determined by their summed probability for equivalence classes of non-deterministic processes. Techniques from metric topology have been used to establish the first result.

We used the cancellativity result in [16] in order to obtain a complete axiomatisation of branching probabilistic bisimilarity. The technical report [15] contains a proof sketch in line with this paper. Yet, as cancellativity is such a fundamental property, and the notion of branching probabilistic bisimulation is mathematically quite involved, we regard it necessary to provide a full, detailed proof.

The continuity results of Section 5, as well as the argumentation from metric topology at other places, are exploited to deal with the uncountable number of inert transitions that arise from combined transitions. One may wonder if the main theorems of the paper can be achieved based on combinatorial arguments. Intuitively, transitions span a convex polyhedron and the uncountability of the branching of transitions may be reduced to the finiteness of the transitions spanning the polyhedron. Despite a number of attempts, we have been forced to leave the question of a simpler combinatorial proof open.

We leave it as open question for future research weather cancellativity holds for larger classes of probabilistic processes, as could be obtained, for instance, by adding recursion, uncountable choice and/or parallel composition to the syntax. A further topic for future research is the study of cancellativity for other weak variants of probabilistic bisimulation, in particular weak probabilistic bisimulation.

Other future work is to be devoted to the construction of an efficient decision algorithm for branching probabilistic bisimilarity. A decision procedure for strong probabilistic bisimilarity based on so-called extended ordered binary trees has been proposed in [4]. An improved algorithm based on partition refinement is presented in [18]. Partition refinement algorithms for weak and branching probabilistic bisimilarity on states are proposed in [29]. Reduction of weak probabilistic bisimilarity checking of the state-based approach of [8] to linear programming is studied in [12]. Although it is currently not clear how to construct an algorithm deciding branching probabilistic bisimilarity as put forward in this paper, it is likely that the procedures of [17] and [29] can serve as a starting point.

References

[1] S. Andova, S. Georgievska & N. Trcka (2012): *Branching bisimulation congruence for probabilistic systems*. *Theoretical Computer Science* 413, pp. 58–72, doi:10.1016/j.tcs.2011.07.020.

[2] S. Andova & T.A.C. Willemse (2006): *Branching bisimulation for probabilistic systems: Characteristics and decidability*. *Theoretical Computer Science* 356, pp. 325–355, doi:10.1016/j.tcs.2006.02.010.

[3] C. Baier, P.R. D'Argenio & H. Hermanns (2020): *On the probabilistic bisimulation spectrum with silent moves*. *Acta Informatica* 57, pp. 465–512, doi:10.1007/s00236-020-00379-2.

[4] C. Baier, B. Engelen & M.E. Majster-Cederbaum (2000): *Deciding bisimilarity and similarity for probabilistic processes*. *Journal of Computer Systems and Sciences* 60(1), pp. 187–231, doi:10.1006/jcss.1999.1683.

[5] C. Baier & M.Z. Kwiatkowska (2000): *Domain equations for probabilistic processes*. *Mathematical Structures in Computer Science* 10(6), pp. 665–717, doi:10.1017/S0960129599002984.

[6] E. Bandini & R. Segala (2001): *Axiomatizations for Probabilistic Bisimulation*. In F. Orejas et al., editor: *Proc. ICALP 2001*, LNCS 2076, pp. 370–381, doi:10.1007/3-540-48224-5_31.

[7] F. Breugel & J. Worrell (2005): *A behavioural pseudometric for probabilistic transition systems*. *Theoretical Computer Science* 331(1), pp. 115–142, doi:10.1016/j.tcs.2004.09.035.

[8] S. Cattani & R. Segala (2002): *Decision Algorithms for Probabilistic Bisimulation*. In L. Brim et al., editor: *Proc. CONCUR 2002*, LNCS 2421, pp. 371–385, doi:10.1007/3-540-45694-5_25.

[9] Y. Deng, R.J. van Glabbeek, M. Hennessy & C.C. Morgan (2009): *Testing Finitary Probabilistic Processes (extended abstract)*. In M. Bravetti & G. Zavattaro, editors: *Proc. CONCUR'09*, LNCS 5710, pp. 274–288, doi:10.1007/978-3-642-04081-8_19.

[10] J. Desharnais, V. Gupta, R. Jagadeesan & P. Panangaden (1999): *Metrics for Labeled Markov Systems*. In J.C.M. Baeten & S. Mauw, editors: *Proc. CONCUR '99*, LNCS 1664, pp. 258–273, doi:10.1007/3-540-48320-9_19.

[11] C. Eisentraut, H. Hermanns, J. Krämer, A. Turrini & L. Zhang (2013): *Deciding Bisimilarities on Distributions*. In K. Joshi et al., editor: *Proc. QEST 2013*, LNCS 8054, pp. 72–88, doi:10.1007/978-3-642-40196-1_6.

[12] L. Ferrer Fioriti, V. Hashemi, H. Hermanns & A. Turrini (2016): *Deciding probabilistic automata weak bisimulation: theory and practice*. *Formal Aspects of Computing* 28(1), pp. 109–143, doi:10.1007/s00165-016-0356-4.

[13] A. Giacalone, Chi-Chang Jou & S.A. Smolka (1990): *Algebraic Reasoning for Probabilistic Concurrent Systems*. In M. Broy & C.B. Jones, editors: *Programming concepts and methods*, North-Holland, pp. 443–458.

[14] M. Giry (1982): *A Categorical Approach to Probability Theory*. In B. Banaschewski, editor: *Categorical Aspects of Topology and Analysis*, LNM 915, pp. 68–85, doi:10.1007/BFb0092872.

[15] R.J. van Glabbeek, J.F. Groote & E.P. de Vink (2019): *A Complete Axiomatization of Branching Bisimilarity for a Simple Process Language with Probabilistic Choice*. Technical Report, Eindhoven University of Technology. Available at `http://rvg.web.cse.unsw.edu.au/pub/AxiomProbBranchingBis.pdf`.

[16] R.J. van Glabbeek, J.F. Groote & E.P. de Vink (2019): *A Complete Axiomatization of Branching Bisimilarity for a Simple Process Language with Probabilistic Choice (extended abstract)*. In M.A. Alvim et al., editor: *The Art of Modelling Computational Systems: A Journey from Logic and Concurrency to Security and Privacy*, LNCS 11760, pp. 139–162, doi:10.1007/978-3-030-31175-9_9.

[17] J.F. Groote & F.W. Vaandrager (1990): *An Efficient Algorithm for Branching Bisimulation and Stuttering Equivalence*. In M. Paterson, editor: *Proc. ICALP '90*, LNCS 443, pp. 626–638, doi:10.1007/BFb0032063.

[18] J.F. Groote, H.J. Rivera Verduzco & E.P. de Vink (2018): *An efficient algorithm to determine probabilistic bisimulation*. *Algorithms* 11(9), pp. 131,1–22, doi:10.3390/a11090131.

[19] J.F. Groote & E.P. de Vink (2019): *An Axiomatization of Strong Distribution Bisimulation for a Language with a Parallel Operator and Probabilistic Choice*. In M.H. ter Beek, A. Fantechi & L. Semini, editors: *From*

Software Engineering to Formal Methods and Tools, and Back, LNCS 11865, pp. 449–463, doi:10.1007/978-3-030-30985-5_26.

[20] H. Hansson & B. Jonsson (1990): *A Calculus for Communicating Systems with Time and Probabilities*. In: *Proc. RTSS 1990*, IEEE, pp. 278–287, doi:10.1109/REAL.1990.128759.

[21] J.I. den Hartog, E.P. de Vink & J.W. de Bakker (2000): *Metric semantics and full abstractness for action refinement and probabilistic choice*. *Electronic Notes in Theoretical Computer Science* 40, pp. 72–99, doi:10.1016/S1571-0661(05)80038-6.

[22] M. Hennessy (2012): *Exploring probabilistic bisimulations, part I*. *Formal Aspects of Computing* 24, pp. 749–768, doi:10.1007/s00165-012-0242-7.

[23] S. Lang (1997): *Undergraduate Analysis (2nd ed.)*. Undergraduate Texts in Mathmatics, Springer, doi:10.1007/978-1-4757-2698-5.

[24] M.D. Lee & E.P. de Vink (2016): *Logical Characterization of Bisimulation for Transition Relations over Probability Distributions with Internal Actions*. In P. Faliszewski, A. Muscholl & R. Niedermeier, editors: *Proc. MFCS 2016*, LIPIcs 58, pp. 29:1–29:14, doi:10.4230/LIPIcs.MFCS.2016.29.

[25] G.J. Norman (1997): *Metric Semantics for Probabilistic Systems*. Ph.D. thesis, Universith of Birmingham.

[26] R. Segala (1995): *Modeling and Verification of Randomized Distributed Real-Time Systems*. Ph.D. thesis, MIT. Technical Report MIT/LCS/TR–676.

[27] R. Segala & N.A. Lynch (1994): *Probabilistic simulations for probabilistic processes*. In B. Jonsson & J. Parrow, editors: *Proc. CONCUR 94*, LNCS 836, pp. 481–496, doi:10.1007/978-3-540-48654-1_35.

[28] M. Stoelinga (2002): Alea Jacta est*: Verification of probabilistic, real-time and parametric systems*. Ph.D. thesis, Radboud Universiteit.

[29] A. Turrini & H. Hermanns (2015): *Polynomial time decision algorithms for probabilistic automata*. *Information and Computation* 244, pp. 134–171, doi:10.1016/j.ic.2015.07.004.