Logical Characterization of Trace Metrics

Valentina Castiglioni
University of Insubria (IT)
v.castiglioni2@uninsubria.it

Simone Tini
University of Insubria (IT)
simone.tini@uninsubria.it

In this paper we continue our research line on logical characterizations of behavioral metrics obtained from the definition of a metric over the set of logical properties of interest. This time we provide a characterization of both strong and weak trace metric on nondeterministic probabilistic processes, based on a minimal boolean logic $\mathbb{L}$ which we prove to be powerful enough to characterize strong and weak probabilistic trace equivalence. Moreover, we also prove that our characterization approach can be restated in terms of a more classic probabilistic $\mathbb{L}$-model checking problem.

1 Introduction

Behavioral equivalences and modal logics have been successfully employed for the specification and verification of communicating concurrent systems, henceforth processes. The former ones provide a simple and elegant tool for comparing the observable behavior of processes. The latter ones allow for an immediate expression of the desired properties of processes. Since the work of [19] on the Hennessy-Milner logic (HML), these two approaches are connected by means of *logical characterizations* of behavioral equivalences: two processes are behaviorally equivalent if and only if they satisfy the same formulae in the logic. Hence, the characterization of an equivalence subsumes both the fact that the logic is as expressive as the equivalence and the fact that the equivalence preserves the logical properties of processes.

It is common agreement that when also quantitative properties of processes are taken into account a metric semantics is favored over behavioral equivalences, [1, 3, 7, 10, 12, 15–17, 23, 24]. since the latter ones are too sensible to small variations in the probabilistic properties of processes. Therefore, the interest in logical characterizations of the so called *behavioral metrics*, namely the quantitative analogues of equivalences that quantify how far the behavior of two processes is apart, is constantly growing.

In this paper we propose a logical characterization of the *strong* and *weak* variants of the *trace metric* [28] for nondeterministic probabilistic processes (PTSs [27]). To this aim we follow the approach of [8] in which a logical characterization of the bisimilarity metric is provided. We introduce two boolean logics $\mathbb{L}$ and $\mathbb{L}_{\mathrm{w}}$, providing a probabilistic choice operator capturing the probability weights that a process assigns to arbitrary traces, which we prove to characterize resp. the *strong* and *weak probabilistic trace equivalences* of [26]. Such a characterization is obtained by introducing the novel notion of *mimicking formulae of resolutions*, i.e. formulae capturing, for each possible resolution of nondeterminism for a process, all the executable traces as well as the probability weights assigned to them. Then we introduce the notions of *distance between formulae* in $\mathbb{L}$ and $\mathbb{L}_{\mathrm{w}}$ which are 1-bounded (pseudo)metrics assigning to each pair of formulae a suitable quantitative analogue of their syntactic disparities. These lift to metrics over processes, called resp. $\mathbb{L}$-*distance* and $\mathbb{L}_{\mathrm{w}}$-*distance*, corresponding to the Hausdorff lifting of the distance between formulae to the sets of formulae satisfied by the two processes. We prove that our $\mathbb{L}$-distance and $\mathbb{L}_{\mathrm{w}}$-distance correspond resp. to the strong and weak trace metric.

An important feature of our characterization method is that, although it is firmly based on the mimicking formulae of resolutions, it does not actually depend on how these resolutions of nondeterminism are

To appear in EPTCS.

obtained from processes. For instance, in this paper we consider resolutions obtained via a *deterministic scheduler* [4,26], but our approach would not be different when applied to *randomized resolutions* [4,26].

Our approach differs from the ones proposed in the literature in that in general logics equipped with a real-valued semantics are used for the characterization, which is then expressed as

$$d(s,t) = \sup_{\varphi \in L} |[\varphi](s) - [\varphi](t)| \tag{1}$$

where d is the behavioral metric of interest, L is the considered logic and $[\varphi](s)$ denotes the value of the formula φ at process s accordingly to the real-valued semantics [1,2,12–14]. In [3] it is proved that the trace metric on Markov Chains (MCs) can be characterized in terms of the probabilistic LTL-model checking problem. Roughly speaking, a characterization as in (1) is obtained from the boolean logic LTL by assigning a real-valued semantics to it, defined by exploiting the probabilistic properties of the MC: the value of a formula $\varphi \in \mathrm{LTL}$ at state s is given by the probability of s to execute a run satisfying φ. In this paper we show that we can obtain a similar result by means of our distance between formulae. More precisely, we provide an alternative characterization of the trace metric on PTSs $\mathbf{d}_T$ in terms of the probabilistic $\mathbb{L}$-model checking problem. In detail, we define a real-valued semantics for $\mathbb{L}$ by assigning to each formula $\Psi \in \mathbb{L}$ at process s the value $[\Psi](s)$ corresponding to the minimal distance between Ψ and any formula satisfied by s. Thus we could use this real-valued semantics to verify whether process s behaves within an allowed tolerance wrt. to the specification given by the formula Ψ. Then, by exploiting some properties of the Hausdorff metric, we will be able to conclude that $\mathbf{d}_T(s,t) = \sup_{\Psi \in \mathbb{L}} |\, [\Psi](s) - [\Psi](t) \,|$ thus giving that the verification of any $\mathbb{L}$-formula in s cannot differ from its verification in t for more than $\mathbf{d}_T(s,t)$ which, in turn, constitutes the maximal observable error in the approximation of s with t.

We can summarize our contributions as follows:

1. Logical characterization of both strong and weak trace metric: we define a distance on the class of formulae $\mathbb{L}$ (resp. $\mathbb{L}_{\mathrm{w}}$) and we prove that the strong (resp. weak) trace metric between two processes equals the syntactic distance between the sets of formulae satisfied by them.
2. Logical characterization of strong trace metric in terms of a probabilistic $\mathbb{L}$-model checking problem: by means of the distance between formulae we equip $\mathbb{L}$ with a real-valued semantics and we use it to establish a characterization of the trace metric as in (1).
3. Logical characterization of both strong and weak probabilistic trace equivalence: by exploiting the notion of mimicking formula, we prove that two processes are strong (resp. weak) trace equivalent if and only if they satisfy the same (resp. syntactically equivalent) formulae in $\mathbb{L}$ (resp. $\mathbb{L}_{\mathrm{w}}$).

2 Background

2.1 Nondeterministic probabilistic transition systems

Nondeterministic probabilistic transition systems [27] combine LTSs [22] and discrete time Markov chains [18,29], allowing us to model reactive behavior, nondeterminism and probability.

As state space we take a set $\mathbf{S}$, whose elements are called *processes*. We let $s,t,\dots$ range over $\mathbf{S}$. Probability distributions over $\mathbf{S}$ are mappings $\pi\colon \mathbf{S} \to [0,1]$ with $\sum_{s\in\mathbf{S}} \pi(s) = 1$ that assign to each $s \in \mathbf{S}$ its probability $\pi(s)$. By $\Delta(\mathbf{S})$ we denote the set of all distributions over $\mathbf{S}$. We let $\pi,\pi',\dots$ range over $\Delta(\mathbf{S})$. For $\pi \in \Delta(\mathbf{S})$, we denote by $\mathsf{supp}(\pi)$ the support of π, namely $\mathsf{supp}(\pi) = \{s \in \mathbf{S} \mid \pi(s) > 0\}$. We consider only distributions with *finite* support. For $s \in \mathbf{S}$ we denote by δ_s the *Dirac distribution* defined by $\delta_s(s) = 1$ and $\delta_s(t) = 0$ for $s \neq t$. The convex combination $\sum_{i\in I} p_i\pi_i$ of a family $\{\pi_i\}_{i\in I}$ of distributions $\pi_i \in \Delta(\mathbf{S})$ with $p_i \in (0,1]$ and $\sum_{i\in I} p_i = 1$ is defined by $(\sum_{i\in I} p_i\pi_i)(s) = \sum_{i\in I}(p_i\pi_i(s))$ for all $s \in \mathbf{S}$.

Definition 1 (PTS, [27])**.** A *nondeterministic probabilistic labeled transition system (PTS)* is a triple $(\mathbf{S}, \mathscr{A}, \rightarrow)$, where: (i) $\mathbf{S}$ is a countable set of processes, (ii) $\mathscr{A}$ is a countable set of *actions*, and (iii) $\rightarrow \subseteq \mathbf{S} \times \mathscr{A} \times \Delta(\mathbf{S})$ is a *transition relation*.

We call $(s,a,\pi) \in \rightarrow$ a *transition*, and we write $s \xrightarrow{a} \pi$ for $(s,a,\pi) \in \rightarrow$. We write $s \xrightarrow{a}$ if there is a distribution $\pi \in \Delta(\mathbf{S})$ with $s \xrightarrow{a} \pi$, and $s \not\xrightarrow{a}$ otherwise. Let $\mathrm{init}(s) = \{a \in \mathscr{A} \mid s \xrightarrow{a}\}$ denote the set of the actions that can be performed by s. Let $\mathrm{der}(s,a) = \{\pi \in \Delta(\mathbf{S}) \mid s \xrightarrow{a} \pi\}$ denote the set of the distributions reachable from s through action a. Let $\mathrm{dpt}(s)$ denote the *depth* of s, namely the maximal number of sequenced transitions that can be performed from s, defined by $\mathrm{dpt}(s) = 0$, if $\mathrm{init}(s) = \emptyset$, and $\mathrm{dpt}(s) = 1 + \sup_{a\in\mathrm{init}(s),\pi\in\mathrm{der}(s,a),t\in\mathsf{supp}(\pi)} \mathrm{dpt}(t)$, otherwise. We say that a process $s \in \mathbf{S}$ is *image-finite* if for all actions $a \in \mathrm{init}(s)$ the set $\mathrm{der}(s,a)$ is finite [20], and that s has *finite depth* if $\mathrm{dpt}(s)$ is finite. Finally, we denote as *finite* the image-finite processes with finite depth. In this paper we consider only processes that are finite.

Throughout the paper we will introduce some equivalence relations on traces and on modal formulae. To deal with the equivalence of probability distributions over these elements, we need to introduce the notion of *lifting* of a relation.

Definition 2. Let X be any set. Consider a relation $\mathscr{R} \subseteq X \times X$. Then the *lifting* of $\mathscr{R}$ is the relation $\mathscr{R}^\dagger \subseteq \Delta(X) \times \Delta(X)$ with $\pi \mathscr{R}^\dagger \pi'$ if whenever $\pi = \sum_{i\in I} p_i \delta_{x_i}$ then $\pi' = \sum_{i\in I, j_i\in J_i} p_{j_i}\delta_{y_{j_i}}$ with $\sum_{j_i\in J_i} p_{j_i} = p_i$ and $x_i \mathscr{R} y_{j_i}$ for all $j_i \in J_i$.

Moreover, we can lift relations to relations over sets. Given a relation $\mathscr{R} \subseteq X \times Y$, we say that two subsets $X' \subseteq X, Y' \subseteq Y$ are in relation $\mathscr{R}$, notation $X' \mathscr{R} Y'$, iff (i) for each $x \in X'$ there is an $y \in Y'$ with $x \mathscr{R} y$, and (ii) for each $y \in Y'$ there is an $x \in X'$ with $x \mathscr{R} y$.

2.2 Strong probabilistic trace equivalence

A probabilistic trace equivalence is a relation over $\mathbf{S}$ that equates processes $s,t \in \mathbf{S}$ if for all resolutions of nondeterminism they can mimic each other's sequences of transitions with the same probability.

Definition 3 (Computation, [4])**.** Let $P = (\mathbf{S}, \mathscr{A}, \rightarrow)$ be a PTS and $s, s' \in \mathbf{S}$. We say that $c := s_0 \xrightarrow{a_1}\!\!\!\!\rightarrow s_1 \xrightarrow{a_2}\!\!\!\!\rightarrow s_2 \dots s_{n-1} \xrightarrow{a_n}\!\!\!\!\rightarrow s_n$ is a *computation* of P of length n from $s = s_0$ to $s' = s_n$ iff for all $i = 1, \dots, n$ there exists a transition $s_{i-1} \xrightarrow{a_i} \pi_i$ in P such that $s_i \in \mathsf{supp}(\pi_i)$, with $\pi_i(s_i)$ being the *execution probability* of step $s_{i-1} \xrightarrow{a_i}\!\!\!\!\rightarrow s_i$ conditioned on the selection of transition $s_{i-1} \xrightarrow{a_i} \pi_i$ of P at s_{i-1}. We denote by $\Pr(c) = \prod_{i=1}^n \pi_i(s_i)$ the product of the execution probabilities of the steps in c.

Let $s, s', s'' \in \mathbf{S}$. Given any computation $c' = s' \xrightarrow{a_1}\!\!\!\!\rightarrow s_1 \xrightarrow{a_2}\!\!\!\!\rightarrow \dots \xrightarrow{a_n}\!\!\!\!\rightarrow s''$ from s' to s'', we write $c = s \xrightarrow{a}\!\!\!\!\rightarrow c'$ if $c = s \xrightarrow{a}\!\!\!\!\rightarrow s' \xrightarrow{a_1}\!\!\!\!\rightarrow \dots \xrightarrow{a_n}\!\!\!\!\rightarrow s''$ is a computation from s to s''.

Given a process $s \in \mathbf{S}$, we say that c is a computation from s if there is a process s' such that c is a computation from s to s'. We denote by $\mathscr{C}(s)$ the set of computations from s. We say that a computation c from process s is *maximal* if it is not a proper prefix of any other computation from s. We denote by $\mathscr{C}_{\max}(s) \subseteq \mathscr{C}(s)$ the subset of the maximal computations from s. Since we are considering finite processes, all computations are guaranteed to be of finite length.

We denote by $\mathscr{A}^\star$ the set of sequences of actions in $\mathscr{A}$ and we call *trace* any element $\alpha \in \mathscr{A}^\star$. We use the special symbol $\varepsilon \notin \mathscr{A}$ to denote the empty trace. We say that a computation is *compatible* with the trace $\alpha \in \mathscr{A}^\star$ iff the sequence of actions labeling the computation steps is equal to α. We

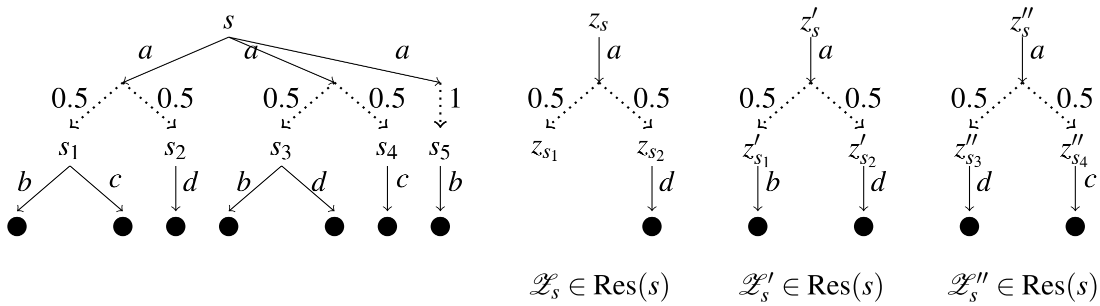

Figure 1: An example of three distinct resolutions for process s. Black circles stand for the probability distribution δ_{nil}, with nil process that cannot execute any action.

denote by $\mathscr{C}(s,\alpha) \subseteq \mathscr{C}(s)$ the set of computations of process s which are compatible with trace α and by $\mathscr{C}_{\max}(s,\alpha) \subseteq \mathscr{C}(s,\alpha)$ we denote the set of maximal computations of s that are compatible with α. Then, given any $\mathscr{C} \subseteq \mathscr{C}(s)$, we define $\Pr(\mathscr{C}) = \sum_{c\in\mathscr{C}} \Pr(c)$.

Definition 4. Let $s \in \mathbf{S}$ and consider any $c \in \mathscr{C}(s)$. We denote by $\mathrm{Tr}(c) \in \mathscr{A}^\star$ the trace to which c is compatible. We extend this notion to sets by letting $\mathrm{Tr}(\mathscr{C}') = \{\mathrm{Tr}(c) \mid c \in \mathscr{C}'\}$ for any $\mathscr{C}' \subseteq \mathscr{C}(s)$. We say that $\mathrm{Tr}(\mathscr{C}(s))$ is the *set of traces* of s and $\mathrm{Tr}(\mathscr{C}_{\max}(s))$ is the *set of maximal traces* of s.

To establish trace equivalence we need first to deal with nondeterministic choices of processes. To this aim, we consider all possible resolutions of nondeterminism one by one. Using the notation of [4], our resolutions correspond to the resolutions obtained via a *deterministic scheduler* (see Fig. 1 for an example).

Definition 5 (Resolution, [4])**.** Let $P = (\mathbf{S},\mathscr{A},\to)$ be a PTS and $s \in \mathbf{S}$. We say that a PTS $\mathscr{Z} = (Z,\mathscr{A},\to_{\mathscr{Z}})$ is a *resolution* for s iff there exists a state correspondence function $\mathrm{corr}_{\mathscr{Z}} \colon Z \to \mathbf{S}$ such that $s = \mathrm{corr}_{\mathscr{Z}}(z_s)$ for some $z_s \in Z$, called the *initial state* of $\mathscr{Z}$, and moreover it holds that:

- $z_s \notin \mathsf{supp}(\pi)$ for any $\pi \in \bigcup_{z\in Z, a\in\mathscr{A}} \mathrm{der}(z,a)$.
- Each $z \in Z \setminus \{z_s\}$ is such that $z \in \mathsf{supp}(\pi)$ for some $\pi \in \bigcup_{z'\in Z\setminus\{z\}, a\in\mathscr{A}} \mathrm{der}(z',a)$.
- Whenever $z \xrightarrow{a}_{\mathscr{Z}} \pi$, then $\mathrm{corr}_{\mathscr{Z}}(z) \xrightarrow{a} \pi'$ with $\pi(z') = \pi'(\mathrm{corr}_{\mathscr{Z}}(z'))$ for all $z' \in Z$.
- Whenever $z \xrightarrow{a_1}_{\mathscr{Z}} \pi_1$ and $z \xrightarrow{a_2}_{\mathscr{Z}} \pi_2$ then $a_1 = a_2$ and $\pi_1 = \pi_2$.

We let $\mathrm{Res}(s)$ be the set of resolutions for s and $\mathrm{Res}(\mathbf{S}) = \bigcup_{s\in\mathbf{S}} \mathrm{Res}(s)$ be the set of all resolutions on $\mathbf{S}$.

Strong probabilistic trace equivalence equates two processes if their resolutions can be matched so that they assign the same probability to all traces.

Definition 6 (Strong probabilistic trace equivalence, [4, 26])**.** Let $P = (\mathbf{S},\mathscr{A},\to)$ be a PTS. We say that $s,t \in \mathbf{S}$ are *strong probabilistic trace equivalent*, notation $s \approx_{\mathrm{st}} t$, iff it holds that:

- For each resolution $\mathscr{Z}_s \in \mathrm{Res}(s)$ of s there is a resolution $\mathscr{Z}_t \in \mathrm{Res}(t)$ of t such that for all traces $\alpha \in \mathscr{A}^\star$ we have $\Pr(\mathscr{C}(z_s,\alpha)) = \Pr(\mathscr{C}(z_t,\alpha))$.
- For each resolution $\mathscr{Z}_t \in \mathrm{Res}(t)$ of t there is a resolution $\mathscr{Z}_s \in \mathrm{Res}(s)$ of s such that for all traces $\alpha \in \mathscr{A}^\star$ we have $\Pr(\mathscr{C}(z_t,\alpha)) = \Pr(\mathscr{C}(z_s,\alpha))$.

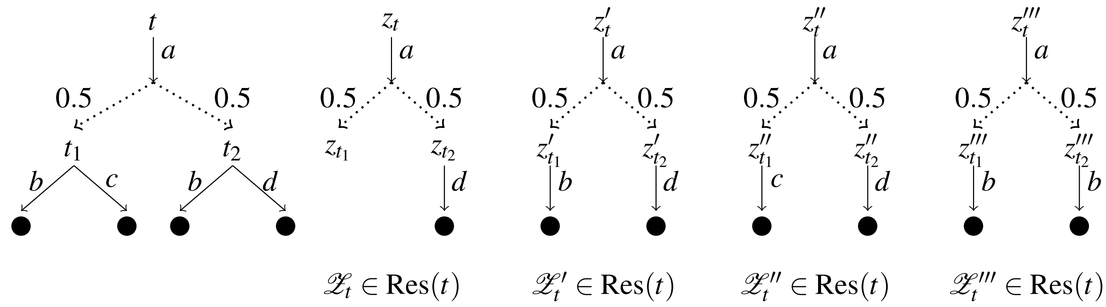

Figure 2: Process t is strong trace equivalent to process s in Fig. 1

***Example* 1.** Consider process s in Fig. 1 and process t in Fig. 2. We have that $s \approx_{\mathrm{st}} t$. Briefly, it is immediate to check that the three resolutions $\mathscr{Z}_s, \mathscr{Z}'_s, \mathscr{Z}''_s \in \mathrm{Res}(s)$ in Fig. 1 are matched resp. by the three resolutions $\mathscr{Z}_t, \mathscr{Z}'_t, \mathscr{Z}''_t \in \mathrm{Res}(t)$ in Fig. 2. Moreover, for all other resolutions, we notice that accordingly to the chosen resolutions for processes t_1 and t_2, process s can always match their traces and related probabilities by selecting the proper a-branch. In particular, resolution $\mathscr{Z}'''_t \in \mathrm{Res}(t)$ in Fig. 2 is matched by the resolution for s corresponding to the rightmost a-branch.

2.3 Weak probabilistic trace equivalence

We extend the set of actions $\mathscr{A}$ to the set $\mathscr{A}_\tau$ containing also the silent action τ. We let $\mathfrak{a}$ range over $\mathscr{A}_\tau$.

Usually, traces are not distinguished by any occurrence of τ in them [28]. Hence, we introduce the notion of *equivalence of traces*.

Definition 7 (Equivalence of traces)**.** The relation of *equivalence of traces* $\equiv_{\mathrm{w}} \subseteq \mathscr{A}_\tau^\star \times \mathscr{A}_\tau^\star$ is the smallest equivalence relation satisfying 1. $\varepsilon \equiv_{\mathrm{w}} \varepsilon$ and 2. given $\alpha = \mathfrak{a}_1\alpha'$, $\beta = \mathfrak{a}_2\beta'$ we have $\alpha \equiv_{\mathrm{w}} \beta$ iff

- either $\mathfrak{a}_1 = \tau$ and $\alpha' \equiv_{\mathrm{w}} \beta$,
- or $\mathfrak{a}_2 = \tau$ and $\alpha \equiv_{\mathrm{w}} \beta'$
- or $\mathfrak{a}_1 = \mathfrak{a}_2$ and $\alpha' \equiv_{\mathrm{w}} \beta'$.

For each trace $\alpha \in \mathscr{A}_\tau^\star$, we denote by $[\alpha]_{\mathrm{w}}$ the equivalence class of α with respect to $\equiv_{\mathrm{w}}$, namely $[\alpha]_{\mathrm{w}} = \{\beta \in \mathscr{A}_\tau^\star \mid \beta \equiv_{\mathrm{w}} \alpha\}$. Moreover, for each computation c, we let $\mathrm{Tr}_{\mathrm{w}}(c) = [\mathrm{Tr}(c)]_{\mathrm{w}}$.

Given any process $s \in \mathbf{S}$ and any trace $\alpha \in \mathscr{A}_\tau^\star$, we say that a computation $c \in \mathscr{C}(s)$ is in $\mathscr{C}^{\mathrm{w}}(s,\alpha)$ iff $\mathrm{Tr}(c) \equiv_{\mathrm{w}} \alpha$ and c is not a proper prefix of any other computation in $\mathscr{C}^{\mathrm{w}}(s,\alpha)$. This is to avoid to count multiple times the same execution probabilities in the evaluation of $\mathrm{Pr}(\mathscr{C}^{\mathrm{w}}(s,\alpha))$.

Definition 8 (Weak probabilistic trace equivalence)**.** Let $P = (\mathbf{S}, \mathscr{A}, \rightarrow)$ be a PTS. We say that $s,t \in \mathbf{S}$ are *weak probabilistic trace equivalent*, notation $s \approx_{\mathrm{wt}} t$, iff it holds that:

- For each resolution $\mathscr{Z}_s \in \mathrm{Res}(s)$ of s there is a resolution $\mathscr{Z}_t \in \mathrm{Res}(t)$ of t such that for all traces $\alpha \in \mathscr{A}^\star$ we have $\mathrm{Pr}(\mathscr{C}^{\mathrm{w}}(z_s,\alpha)) = \mathrm{Pr}(\mathscr{C}^{\mathrm{w}}(z_t,\alpha))$.
- For each resolution $\mathscr{Z}_t \in \mathrm{Res}(t)$ of t there is a resolution $\mathscr{Z}_s \in \mathrm{Res}(s)$ of s such that for all traces $\alpha \in \mathscr{A}^\star$ we have $\mathrm{Pr}(\mathscr{C}^{\mathrm{w}}(z_t,\alpha)) = \mathrm{Pr}(\mathscr{C}^{\mathrm{w}}(z_s,\alpha))$.

3 Trace metrics

In this section we introduce the quantitative analogues of strong and weak probabilistic trace equivalence, namely the *strong* and *weak trace metric*, resp., which are 1-bounded pseudometrics that quantify how much the behavior of two processes is apart wrt. the strong (resp. weak) probabilistic trace semantics. Our metrics are a revised version of the trace metric proposed in [28]. Briefly, in [28] there is a distinction between the notions of *path* and *trace*: any $\alpha \in \mathscr{A}_\tau^\star$ is called path and the trace related to a path is obtained by deleting any occurrence of τ from it. The metric in [28] is then defined only on traces and it has inspired our strong trace metric. In the present paper we distinguish between the strong and the weak case and we regain the results in [28] by our equivalence of traces: the weak trace metric coincides with the strong one on the quotient space wrt. $\equiv_{\mathrm{w}}$.

3.1 The Kantorovich and Hausdorff lifting functionals

In the literature we can find several examples of behavioral metrics on systems with probability and nondeterminism (see among others [1,5,6,10,12,28]). In this paper we follow the approach of [6,10,28] in which two kind of metrics are combined to obtain a metric on the system. The *Kantorovich metric* [21] quantifies the disparity between the probabilistic properties of processes and it is defined by means of the notion of *matching*. For any set X, a matching for distributions $\pi,\pi' \in \Delta(X)$ is a distribution over the product space $\mathfrak{w} \in \Delta(X \times X)$ with π and π' as left and right marginal resp., namely $\sum_{y \in X} \mathfrak{w}(x,y) = \pi(x)$ and $\sum_{x \in X} \mathfrak{w}(x,y) = \pi'(y)$ for all $x,y \in X$. Let $\mathfrak{W}(\pi,\pi')$ denote the set of all matchings for π,π'.

Definition 9 (Kantorovich metric, [21])**.** Let $d\colon X \times X \to [0,1]$ be a 1-bounded metric. The *Kantorovich lifting* of d is the 1-bounded metric $\mathbf{K}(d)\colon \Delta(X) \times \Delta(X) \to [0,1]$ defined for all $\pi,\pi' \in \Delta(X)$ by

$$\mathbf{K}(d)(\pi,\pi') = \min_{\mathfrak{w} \in \mathfrak{W}(\pi,\pi')} \sum_{x,y \in X} \mathfrak{w}(x,y) \cdot d(x,y).$$

We remark that since we are considering only probability distributions with finite support, the minimum over $\mathfrak{W}(\pi,\pi')$ is well defined for all $\pi,\pi' \in \Delta(X)$.

The *Hausdorff metric* allows us to lift any distance over probability distributions to a distance over sets of probability distributions.

Definition 10 (Hausdorff metric)**.** Let $\hat{d}\colon \Delta(X) \times \Delta(X) \to [0,1]$ be a 1-bounded metric. The *Hausdorff lifting* of $\hat{d}$ is the 1-bounded metric $\mathbf{H}(\hat{d})\colon \mathscr{P}(\Delta(X)) \times \mathscr{P}(\Delta(X)) \to [0,1]$ defined by

$$\mathbf{H}(\hat{d})(\Pi_1,\Pi_2) = \max\left\{ \sup_{\pi_1 \in \Pi_1} \inf_{\pi_2 \in \Pi_2} \hat{d}(\pi_1,\pi_2), \sup_{\pi_2 \in \Pi_2} \inf_{\pi_1 \in \Pi_1} \hat{d}(\pi_2,\pi_1) \right\}$$

for all $\Pi_1,\Pi_2 \subseteq \Delta(X)$, where $\inf \emptyset = 1$, $\sup \emptyset = 0$.

Hence, given two processes $s,t \in \mathbf{S}$, the idea is to quantify the distance between each pair of their resolutions by exploiting the Kantorovich metric, which quantifies the disparities in the probabilities of the two processes to execute the same traces. Then, we lift this distance on resolutions to a distance between s and t by means of the Hausdorff metric. Intuitively, as each resolution captures a different set of nondeterministic choices of a process, we use the Hausdorff metric to compare the possible choices of the two processes and to match them in order to obtain the minimal distance.

3.2 Strong trace metric

To define the strong trace metric we start from a distance between traces, defined as the discrete metric over traces: two traces are at distance 1 if they are distinct, otherwise the distance is set to 0. Differently from [28] we do not consider any discount on the distance between traces. Trace equivalences, and thus metrics, are usually employed when the observations on the system cannot be done in a step-by-step fashion, but only the total behavior of the system can be observed. Hence, a step-wise discount does not fit in this setting. However, the discount would not introduce any technical issue.

Definition 11 (Distance between traces)**.** The *distance between traces* $d_T : \mathscr{A}^\star \times \mathscr{A}^\star \to [0,1]$ is defined for any pair of traces $\alpha, \beta \in \mathscr{A}^\star$ by

$$d_T(\alpha,\beta) = \begin{cases} 0 & \text{if } \alpha = \beta \\ 1 & \text{otherwise.} \end{cases}$$

Following [28] we aim to lift the distance d_T to a distance between resolutions by means of the Kantorovich lifting functionalwhich, we recall, is defined on probability distributions. As shown in the following example, we are not guaranteed that the function $\mathrm{Pr}(\mathscr{C}(\_,\_))$ defines a probability distribution on the set of traces of a resolution.

***Example* 2.** Consider process t and the resolution $\mathscr{Z}_r \in \mathrm{Res}(t)$ for it, represented in Fig. 2. We can distinguish three computations for z_t:

$$\begin{aligned} c_1 &= z_t \xrightarrow{a}\!\!\!\!\rightarrow z_{t_1} \\ c_2 &= z_t \xrightarrow{a}\!\!\!\!\rightarrow z_{t_2} \\ c_3 &= z_t \xrightarrow{a}\!\!\!\!\rightarrow z_{t_2} \xrightarrow{d}\!\!\!\!\rightarrow \text{nil}. \end{aligned}$$

Clearly, $\mathrm{Tr}(\mathscr{C}(z_t)) = \{a, ad\}$. Then we have

$$\begin{aligned} \mathrm{Pr}(\mathscr{C}(z_t,a)) &= \textstyle\sum_{c\in\mathscr{C}(z_t,a)} \mathrm{Pr}(c) = \mathrm{Pr}(c_1) + \mathrm{Pr}(c_2) = 1 \\ \mathrm{Pr}(\mathscr{C}(z_t,ad)) &= \textstyle\sum_{c\in\mathscr{C}(z_t,ad)} \mathrm{Pr}(c) = \mathrm{Pr}(c_3) = 0.5 \end{aligned}$$

from which we gather

$$\sum_{\alpha\in\mathrm{Tr}(\mathscr{C}(z_t))} \mathrm{Pr}(\mathscr{C}(z_t,\alpha)) = \mathrm{Pr}(\mathscr{C}(z_t,a)) + \mathrm{Pr}(\mathscr{C}(z_t,ad)) = 1 + 0.5 > 1.$$

However, as shown in the following lemma, if we consider only maximal computations we obtain a probability distribution over traces.

Lemma 1. *Consider any resolution $\mathscr{Z} \in \mathrm{Res}(\mathbf{S})$ with initial state z. We have that $\sum_{c\in\mathscr{C}_{\max}(z)} \mathrm{Pr}(c) = 1$.*

Proof. We proceed by induction over the depth of z.

The base case $\mathrm{dpt}(z) = 0$ is immediate since we have that $\mathscr{C}(z) = \{\varepsilon\}$ and $\mathrm{Pr}(\varepsilon) = 1$.

Consider now the inductive step $\mathrm{dpt}(z) > 0$. Assume, wlog., that $z \xrightarrow{a}_{\mathscr{Z}} \pi$. Therefore, each trace $c \in \mathscr{C}_{\max}(z)$ will be of the form $c = z \xrightarrow{a}\!\!\!\!\rightarrow c'$ for some $c' \in \mathscr{C}_{\max}(z')$ for any $z' \in \mathsf{supp}(\pi)$ and moreover for such a trace c it holds that $\mathrm{Pr}(c) = \pi(z')\mathrm{Pr}(c')$. Thus we have

$$\begin{aligned} \textstyle\sum_{c\in\mathscr{C}_{\max}(z)} \mathrm{Pr}(c) &= \textstyle\sum_{\substack{z'\in\mathsf{supp}(\pi)\\ c'\in\mathscr{C}_{\max}(z')}} \pi(z')\mathrm{Pr}(c') \\ &= \textstyle\sum_{z'\in\mathsf{supp}(\pi)} \pi(z') \Big(\sum_{c'\in\mathscr{C}_{\max}(z')} \mathrm{Pr}(c') \Big) \\ &= \textstyle\sum_{z'\in\mathsf{supp}(\pi)} \pi(z') \cdot 1 && \text{(by induction over } \mathrm{dpt}(z') < \mathrm{dpt}(z)) \\ &= 1. \end{aligned}$$

□

Definition 12 (Trace distribution)**.** Consider any resolution $\mathscr{Z} \in \mathrm{Res}(\mathbf{S})$, with initial state z. We define the *trace distribution* of $\mathscr{Z}$ as the function $\mathscr{T}_{\mathscr{Z}} \colon \mathscr{A}^\star \to [0,1]$ defined for each $\alpha \in \mathscr{A}^\star$ by

$$\mathscr{T}_{\mathscr{Z}}(\alpha) = \mathrm{Pr}(\mathscr{C}_{\max}(z,\alpha)).$$

Notice that only maximal computations are in the support of $\mathscr{T}_{\mathscr{Z}}$. This guarantees that $\mathscr{T}_{\mathscr{Z}}$ is a distribution.

Lemma 2. *Consider any resolution $\mathscr{Z} \in \mathrm{Res}(\mathbf{S})$, with initial state z. Then the trace distribution $\mathscr{T}_{\mathscr{Z}}$ of $\mathscr{Z}$ is a probability distribution over $\mathscr{A}^\star$.*

Proof. By definition and by Lemma 1 we have that for each $\alpha \in \mathscr{A}^\star$

$$0 \le \mathrm{Pr}(\mathscr{C}_{\max}(z,\alpha)) = \sum_{c \in \mathscr{C}_{\max}(z,\alpha)} \mathrm{Pr}(c) \le \sum_{c \in \mathscr{C}_{\max}(z)} \mathrm{Pr}(c) = 1$$

Hence, we are guaranteed that $\mathscr{T}_{\mathscr{Z}}(\alpha) \in [0,1]$ for each $\alpha \in \mathscr{A}^\star$. Thus, to prove the thesis we simply need to show that $\sum_{\alpha \in \mathscr{A}^\star} \mathscr{T}_{\mathscr{Z}}(\alpha) = 1$. We have that

$$\begin{aligned}
\textstyle\sum_{\alpha \in \mathscr{A}^\star} \mathscr{T}_{\mathscr{Z}}(\alpha) = {} & \textstyle\sum_{\alpha \in \mathscr{A}^\star} \mathrm{Pr}(\mathscr{C}_{\max}(z,\alpha)) \\
= {} & \textstyle\sum_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))} \mathrm{Pr}(\mathscr{C}_{\max}(z,\alpha)) \\
= {} & \textstyle\sum_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z)), c \in \mathscr{C}_{\max}(z,\alpha)} \mathrm{Pr}(c) \\
= {} & \textstyle\sum_{c \in \bigcup_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))} \mathscr{C}_{\max}(z,\alpha)} \mathrm{Pr}(c) \\
= {} & \textstyle\sum_{c \in \mathscr{C}_{\max}(z)} \mathrm{Pr}(c) \\
= {} & 1
\end{aligned}$$

where

- the second equality follows from the fact that by definition $\mathrm{Pr}(\mathscr{C}_{\max}(z,\alpha)) = 0$ for each $\alpha \notin \mathrm{Tr}(\mathscr{C}_{\max}(z))$;
- the fourth equality follows from the fact that each maximal computation of z belongs to a set $\mathscr{C}_{\max}(z,\alpha)$ for at most one trace α, namely $\bigcup_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))} \mathscr{C}_{\max}(z,\alpha)$ is a disjoint union (and therefore no probability weight is counted more than once);
- the fifth equality follows by the fact that the disjoint union $\bigcup_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))} \mathscr{C}_{\max}(z,\alpha)$ is a partition of $\mathscr{C}_{\max}(z)$;
- the sixth equality follows by Lemma 1.

□

We remark that function $\mathscr{T}_{\_}$ plays the role of the *trace distribution* introduced in [26]. Formally, in [26] the trace distribution for a resolution is defined as the probability space built over its set of traces. Here, we simply identify it with the probability distribution defined on the probability space. In this setting, two resolutions are said to be *trace distribution equivalent* if they have the same trace distribution and thus two processes are trace equivalent if their resolutions are pairwise equivalent.

Lemma 3. *Consider any resolution $\mathscr{Z} \in \mathrm{Res}(\mathbf{S})$ with initial state z. Consider any trace $\alpha \in \mathscr{A}^\star$. Then $\mathrm{Pr}(\mathscr{C}(z,\alpha)) = \sum_{c \in P_{\max}(z,\alpha)} \mathrm{Pr}(c)$, where $P_{\max}(z,\alpha)$ is the set of maximal computations from z having a prefix which is compatible with α.*

Proof. For simplicity let us distinguish two cases.

1. $\Pr(\mathscr{C}(z,\alpha)) = 0$. This implies that there is no computation from z which is compatible with α. Clearly, this gives that there can not be any maximal computation from z having a prefix compatible with α, namely $P_{\max}(z,\alpha) = \emptyset$. Thus we have $\sum_{c\in P_{\max}(z,\alpha)} \Pr(c) = 0$ from which the thesis follows.
2. $\Pr(\mathscr{C}(z,\alpha)) > 0$. In this case, we proceed by induction over $|\alpha|$.
   - Base case $|\alpha| = 0$, namely $\alpha = \varepsilon$. The only computation compatible with α is the empty computation for which it holds that $\Pr(\mathscr{C}(z,\alpha)) = 1$. Since the empty computation is a prefix for all computations from z we have that $P_{\max}(z,\alpha) = \mathscr{C}_{\max}(z)$. By Lemma 1 we have that $\sum_{c\in\mathscr{C}_{\max}(z)} \Pr(c) = 1$ and thus the thesis follows.
   - Inductive step $|\alpha| > 0$. Assume wlog that the only transition inferable for z in $\mathscr{Z}$ is $z \xrightarrow{a}_{\mathscr{Z}} \pi$. Hence $\alpha = a\alpha'$ for some $\alpha' \in \mathscr{A}^\star$, with $|\alpha'| < |\alpha|$. Then we have

$$\begin{aligned}
\Pr(\mathscr{C}(z,\alpha)) =\; & \textstyle\sum_{z'\in\mathsf{supp}(\pi)} \pi(z')\Pr(\mathscr{C}(z',\alpha')) \\
=\; & \textstyle\sum_{z'\in\mathsf{supp}(\pi)} \left(\pi(z')\cdot \sum_{c'\in P_{\max}(z',\alpha')} \Pr(c')\right) \quad \text{(by induction over } |\alpha'|) \\
=\; & \textstyle\sum_{z'\in\mathsf{supp}(\pi),\, c'\in P_{\max}(z',\alpha')} \pi(z')\Pr(c') \\
=\; & \textstyle\sum_{c\in P_{\max}(z,a\alpha')} \Pr(c)
\end{aligned}$$

     where the last equality follows by considering that

$$P_{\max}(z,a\alpha') = \left\{c \mid c = z \xrightarrow{a}_{\mathscr{Z}} c' \text{ and } c' \in \bigcup_{z'\in\mathsf{supp}(\pi)} P_{\max}(z',\alpha')\right\}.$$

□

Proposition 1. *For any pair of resolutions $\mathscr{Z}_1,\mathscr{Z}_2 \in \mathrm{Res}(\mathbf{S})$, with initial states z_1, z_2 resp., we have that $\mathscr{T}_{\mathscr{Z}_1} = \mathscr{T}_{\mathscr{Z}_2}$ iff $\Pr(\mathscr{C}(z_1,\alpha)) = \Pr(\mathscr{C}(z_2,\alpha))$ for all traces $\alpha \in \mathscr{A}^\star$.*

Proof. The thesis follows by applying the same arguments used it the proof of Theorem 2 below. □

Hence, we can now follow [28] to define the *trace metric*.

Definition 13 (Trace distance between resolutions)**.** The *trace distance between resolutions* $D_T\colon \mathrm{Res}(\mathbf{S})\times\mathrm{Res}(\mathbf{S}) \to [0,1]$ is defined for any $\mathscr{Z}_1,\mathscr{Z}_2 \in \mathrm{Res}(\mathbf{S})$ by

$$D_T(\mathscr{Z}_1,\mathscr{Z}_2) = \mathbf{K}(d_T)(\mathscr{T}_{\mathscr{Z}_1},\mathscr{T}_{\mathscr{Z}_2}).$$

Proposition 2 ([28, Proposition 2])**.** *The kernel of D_T is strong trace distribution equivalence of resolutions.*

To deal with nondeterministic choices, we lift the distance over deterministic resolutions to a pseudometric over processes by means of the Hausdorff lifting functional.

Definition 14 (Strong trace metric)**.** *Strong trace metric* $\mathbf{d}_T\colon \mathbf{S}\times\mathbf{S} \to [0,1]$ is defined for all $s,t \in \mathbf{S}$ as

$$\mathbf{d}_T(s,t) = \mathbf{H}(D_T)(\mathrm{Res}(s),\mathrm{Res}(t)).$$

Proposition 3 ([28, Proposition 3])**.** *The kernel of $\mathbf{d}_T$ is probabilistic strong trace equivalence.*

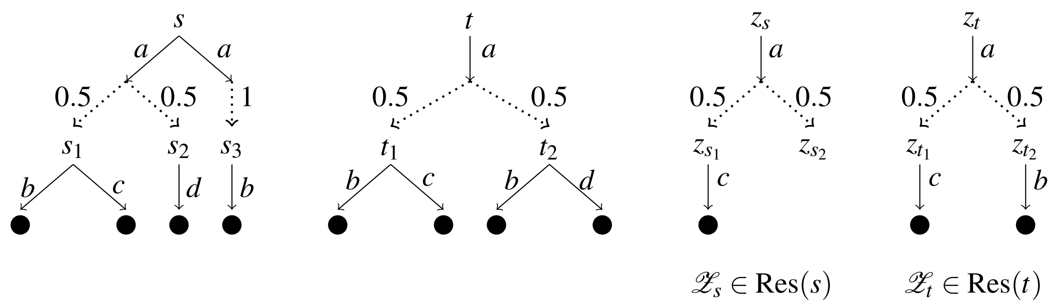

Figure 3: Processes s,t are such that $s \not\approx_{\mathrm{st}} t$ and $\mathbf{d}_T(s,t) = 0.5$.

***Example* 3.** Consider processes s,t in Fig. 3. We have that $s \not\approx_{\mathrm{st}} t$. Notice that none of the resolutions for s can exhibit both traces ab and ac. Thus, whenever we chose resolution $\mathscr{Z}_t \in \mathrm{Res}(t)$ in Fig. 3 for t, then there is no resolution for s that can match $\mathscr{Z}_t$ on all traces.

Let us evaluate the trace distance between s and t. Since resolution $\mathscr{Z}_t$ for t distinguishes the two processes, we start by evaluating its distance from the resolutions for s. Consider the resolution $\mathscr{Z}_s \in \mathrm{Res}(s)$ in Fig. 3. By Def. 12, we have

$$\mathscr{T}_{\mathscr{Z}_s} = 0.5\delta_{ac} + 0.5\delta_a \qquad \mathscr{T}_{\mathscr{Z}_t} = 0.5\delta_{ac} + 0.5\delta_{ab}.$$

Clearly, $d_T(ac,ac) = 0$ and $d_T(ac,a) = d_T(ac,ab) = d_T(a,ab) = 1$. Thus, by Def 13 we have

$$\begin{aligned} D_T(\mathscr{Z}_s, \mathscr{Z}_t) = &\ \mathbf{K}(d_T)(\mathscr{T}_{\mathscr{Z}_s}, \mathscr{T}_{\mathscr{Z}_t}) \\ = &\ \min_{\mathfrak{w} \in \mathfrak{W}(\mathscr{T}_{\mathscr{Z}_s}, \mathscr{T}_{\mathscr{Z}_t})} \textstyle\sum_{\alpha \in \mathsf{supp}(\mathscr{T}_{\mathscr{Z}_s}), \beta \in \mathsf{supp}(\mathscr{T}_{\mathscr{Z}_t})} \mathfrak{w}(\alpha,\beta) \cdot d_T(\alpha,\beta) \\ = &\ 0.5 \cdot d_T(ac,ac) + 0.5 \cdot d_T(a,ab) \\ = &\ 0.5 \end{aligned}$$

where to minimize the distance we have matched the two occurrences of the trace ac. By similar calculations, one can easily obtain that

$$0.5 = D_T(\mathscr{Z}_t, \mathscr{Z}_s) = \sup_{\mathscr{Z}_2 \in \mathrm{Res}(t)} \inf_{\mathscr{Z}_1 \in \mathrm{Res}(s)} D_T(\mathscr{Z}_2, \mathscr{Z}_1).$$

Moreover, it is immediate to check that whichever resolution for s we choose, there is always a resolution for t which is at trace distance 0 from it, namely

$$0 = \sup_{\mathscr{Z}_1 \in \mathrm{Res}(s)} \inf_{\mathscr{Z}_2 \in \mathrm{Res}(t)} D_T(\mathscr{Z}_1, \mathscr{Z}_2).$$

Therefore, we can conclude that

$$\mathbf{d}_T(s,t) = \mathbf{H}(D_T)(\mathrm{Res}(s), \mathrm{Res}(t)) = \max\{0,\ 0.5\} = 0.5$$

3.3 Weak trace metric

To obtain the quantitative analogue of the weak trace equivalence, it is enough to adapt the notion of distance between traces (Definition 11) to the weak context. The idea is that since silent steps cannot be observed, then they should not count on the trace distance. Thus we introduce the notion of *weak distance between traces* which is a 1-bounded pseudometric over $\mathscr{A}_\tau^\star$ having $\equiv_{\mathrm{w}}$ as kernel.

Definition 15 (Weak distance between traces)**.** The *weak distance between traces* $d_T^{\mathrm{w}} \colon \mathscr{A}_\tau^\star \times \mathscr{A}_\tau^\star \to [0,1]$ is defined for any pair of traces $\alpha, \beta \in \mathscr{A}_\tau^\star$ by

$$d_T^{\mathrm{w}}(\alpha,\beta) = \begin{cases} 0 & \text{if } \alpha \equiv_{\mathrm{w}} \beta \\ 1 & \text{otherwise.} \end{cases}$$

It is clear that d_T^{w} is a 1-bounded pseudometric whose kernel is the equivalence of traces.

By substituting d_T with d_T^{w} in Definition 13 we obtain the notion of *weak trace distance between resolutions*, denoted by the 1-bounded pseudometric D_T^{w}. By lifting the relation of equivalence of traces $\equiv_{\mathrm{w}}$ to an equivalence on probability distributions over traces $\equiv_{\mathrm{w}}^{\dagger}$, we obtain that the kernel of D_T^{w} is given by the lifted equivalence on trace distributions, namely by the weak trace distribution equivalence of resolutions. We can prove that our characterization of weak trace equivalence is equivalent to the one proposed in [26] in terms of trace distributions.

To simplify the reasoning in the upcoming proofs, let us define the weak version of the trace distribution given in Definition 12. The idea is that we want to define a probability distribution on the traces executable by a resolution up-to trace equivalence.

Definition 16. Let $s \in \mathbf{S}$ and consider any resolution $\mathscr{Z} \in \mathrm{Res}(\mathbf{S})$, with $z = \mathrm{corr}_{\mathscr{Z}}^{-1}(s)$. We define the *weak trace distribution* for $\mathscr{Z}$ as the function $\mathscr{T}_{\mathscr{Z}}^{\mathrm{w}} \colon \mathscr{A}_\tau^\star \to [0,1]$ defined by $\mathscr{T}_{\mathscr{Z}}^{\mathrm{w}}(\alpha) = \Pr(\mathscr{C}_{\max}^{\mathrm{w}}(z,\alpha))$.

Lemma 4. *For each $\mathscr{Z} \in \mathrm{Res}(\mathbf{S})$, the weak trace distribution $\mathscr{T}_{\mathscr{Z}}^{\mathrm{w}}$ is a probability distribution over $\mathscr{A}^\star$.*

Proof. The thesis follows by applying the same arguments used in the proof of Lemma 2 above. □

Remark 1. Notice that $\mathscr{T}_{-}^{\mathrm{w}}$ is not a probability distribution over $\mathscr{A}_\tau^\star$. In fact it is enough to consider the simple resolution $\mathscr{Z}$ having z as initial state for which the only transition in $\mathscr{Z}$ is $c = z \xrightarrow{a}_{\mathscr{Z}} \delta_{\mathrm{nil}}$, namely z executes a and then with probability 1 it ends its execution. Clearly we have that $a \equiv_{\mathrm{w}} \tau^n a \tau^m$ for all $n,m \geq 0$. Let $\alpha_{n,m} = \tau^n a \tau^m$. Then by definition of weak trace distribution (Definition 16) we would have that $\mathscr{T}_{\mathscr{Z}}^{\mathrm{w}}(\alpha_{n,m}) = \Pr(\mathscr{C}_{\max}^{\mathrm{w}}(z,\alpha_{n,m})) = \Pr(c) = 1$, for all $n,m \geq 0$. Clearly this would imply that $\sum_{\alpha \in \mathscr{A}_\tau^\star} \mathscr{T}_{\mathscr{Z}}^{\mathrm{w}}(\alpha) = \sum_{n,m\geq 0} \mathscr{T}_{\mathscr{Z}}^{\mathrm{w}}(\alpha_{n,m}) > 1$.

However we remark hat $\mathscr{T}_{\mathscr{Z}}$ is a probability distribution over $\mathscr{A}_\tau^\star$ and thus D_T^{w} is well defined.

We aim to show now that there is a strong relation between the trace distribution for a resolution and its weak version: they are equivalent distributions.

Lemma 5. *For each $\mathscr{Z} \in \mathrm{Res}(\mathbf{S})$ we have that $\mathscr{T}_{\mathscr{Z}} \equiv_{\mathrm{w}}^{\dagger} \mathscr{T}_{\mathscr{Z}}^{\mathrm{w}}$.*

Proof. The thesis follows by applying the same arguments used in the proof of Lemma 8 below. □

Proposition 4. *For any pair of resolutions $\mathscr{Z}_1, \mathscr{Z}_2 \in \mathrm{Res}(\mathbf{S})$, with initial states z_1 and z_2 resp., we have that $\mathscr{T}_{\mathscr{Z}_1} \equiv_{\mathrm{w}}^{\dagger} \mathscr{T}_{\mathscr{Z}_2}$ iff $\Pr(\mathscr{C}^{\mathrm{w}}(z_1,\alpha)) = \Pr(\mathscr{C}^{\mathrm{w}}(z_2,\alpha))$ for all $\alpha \in \mathscr{A}^\star$.*

Proof. The thesis follows by the same arguments used in the proof of Theorem 4 below. □

Proposition 5. *The kernel of D_T^{w} is weak trace distribution equivalence of resolutions.*

Proof. The thesis follows by the same arguments used in the proof of Theorem 9 below. □

By substituting D_T with D_T^{w} in Definition 14 we obtain the notion of *weak trace metric*, denoted by the 1-bounded pseudometric $\mathbf{d}_T^{\mathrm{w}}$.

Definition 17 (Weak trace metric)**.** The *weak trace metric* $\mathbf{d}_T^{\mathrm{w}} \colon \mathbf{S} \times \mathbf{S} \to [0,1]$ is defined for all $s,t \in \mathbf{S}$ as

$$\mathbf{d}_T^{\mathrm{w}}(s,t) = \mathbf{H}(D_T^{\mathrm{w}})(\mathrm{Res}(s), \mathrm{Res}(t)).$$

The kernel of the weak trace metric is weak trace equivalence.

Proposition 6. *The kernel of* $\mathbf{d}_T^{\mathrm{w}}$ *is probabilistic weak trace equivalence.*

Proof. ($\Rightarrow$) Assume first that $\mathbf{d}_T^{\mathrm{w}}(s,t) = 0$. We aim to show that $s \approx_{\mathrm{wt}} t$. Since

- by definition $\mathbf{d}_T^{\mathrm{w}}(s,t) = \mathbf{H}(D_T^{\mathrm{w}})(\mathrm{Res}(s), \mathrm{Res}(t))$ and
- the kernel of D_T^{w} is $\equiv_{\mathrm{w}}^{\dagger}$ by Proposition 5

from $\mathbf{d}_T^{\mathrm{w}}(s,t) = 0$ we can infer that $\mathrm{Res}(s) \equiv_{\mathrm{w}}^{\dagger} \mathrm{Res}(t)$. Then, by Proposition 4 we can conclude that $s \approx_{\mathrm{wt}} t$.

($\Leftarrow$) Assume now that $s \approx_{\mathrm{wt}} t$. We aim to show that this implies that $\mathbf{d}_T^{\mathrm{w}}(s,t) = 0$. By Proposition 4 we have that $s \approx_{\mathrm{wt}} t$ implies that $\mathrm{Res}(s) \equiv_{\mathrm{w}}^{\dagger} \mathrm{Res}(t)$. Since the kernel of D_T^{w} is given by $\equiv_{\mathrm{w}}^{\dagger}$ (Proposition 5), we can infer

$$\mathbf{d}_T^{\mathrm{w}}(s,t) = \mathbf{H}(D_T^{\mathrm{w}})(\mathrm{Res}(s), \mathrm{Res}(t)) = 0.$$

□

4 Modal logics for traces

In this section we introduce the modal logics $\mathbb{L}$ and $\mathbb{L}_{\mathrm{w}}$ that will allow us to characterize resp. the strong trace equivalence and its weak version, as well as their quantitative counterparts. The classes $\mathbb{L}$ and $\mathbb{L}_{\mathrm{w}}$ are a simplified version of the modal logic $\mathscr{L}$ [11] which has been successfully employed in [8] to characterize the bisimilarity metric [6, 10, 12].

The logic $\mathbb{L}$ consists of two classes of formulae: the class $\mathbb{L}^{\mathrm{t}}$ of *trace formulae*, which are constituted by (finite) sequences of diamond operators and that will be used to represent traces, and the class $\mathbb{L}^{\mathrm{d}}$ of *trace distribution formulae*, which are defined as probability distributions over trace formulae and that will be used to capture the quantitative properties of resolutions, and thus of processes.

Definition 18 (Modal logic $\mathbb{L}$)**.** The classes of *trace distribution formulae* $\mathbb{L}^{\mathrm{d}}$ and *trace formulae* $\mathbb{L}^{\mathrm{t}}$ over $\mathscr{A}$ are defined by the following BNF-like grammar:

$$\mathbb{L}^{\mathrm{d}}\colon\ \Psi ::= \bigoplus_{i \in I} r_i \Phi_i \qquad\qquad \mathbb{L}^{\mathrm{t}}\colon\ \Phi ::= \top \mid \langle a \rangle \Phi$$

where: (i) Ψ ranges over $\mathbb{L}^{\mathrm{d}}$, (ii) Φ ranges over $\mathbb{L}^{\mathrm{t}}$, (iii) $a \in \mathscr{A}$, (iv) $I \neq \emptyset$ is a finite set of indexes, (v) the formulae Φ_i for $i \in I$ are pairwise distinct, namely $\Phi_i \neq \Phi_j$ for each $i,j \in I$ with $i \neq j$ and (vi) for all $i \in I$ we have $r_i \in (0,1]$ and $\sum_{i \in I} r_i = 1$.

To improve readability, we shall write $r_1\Phi_1 \oplus r_2\Phi$ for $\bigoplus_{i \in I} r_i \Phi_i$ with $I = \{1,2\}$ and Φ for $\bigoplus_{i \in I} r_i \Phi_i$ with $I = \{i\}$, $r_i = 1$ and $\Phi_i = \Phi$.

Definition 19 (Depth)**.** The *depth of trace distribution formulae* in $\mathbb{L}^{\mathrm{d}}$ is defined as $\mathrm{dpt}(\bigoplus_{i \in I} r_i \Phi_i) = \max_{i \in I} \mathrm{dpt}(\Phi_i)$ where the *depth of trace formulae* in $\mathbb{L}^{\mathrm{t}}$ is defined by induction on their structure as (i) $\mathrm{dpt}(\top) = 0$ and (ii) $\mathrm{dpt}(\langle a \rangle \Phi) = 1 + \mathrm{dpt}(\Phi)$.

Definition 20 (Semantics of $\mathbb{L}^{\mathrm{t}}$)**.** The *satisfaction relation* $\models \subseteq \mathscr{C} \times \mathbb{L}^{\mathrm{t}}$ is defined by structural induction over trace formulae in $\mathbb{L}^{\mathrm{t}}$ by

- $c \models \top$ always;
- $c \models \langle a \rangle \Phi$ iff $c = s \xrightarrow{a}\!\!\!\!\rightarrow c'$ for some computation c' such that $c' \models \Phi$.

We say that a computation c from a process s is *compatible* with the trace formula $\Phi \in \mathbb{L}^{\mathrm{t}}$, notation $c \in \mathscr{C}^{\mathrm{t}}(s, \Phi)$, if $c \models \Phi$ and $|c| = \mathrm{dpt}(\Phi)$.

Definition 21 (Semantics of $\mathbb{L}^{\mathrm{d}}$)**.** The *satisfaction relation* $\models \subseteq \mathbf{S} \times \mathbb{L}^{\mathrm{d}}$ is defined by

- $s \models \bigoplus_{i \in I} r_i \Phi_i$ iff there is a resolution $\mathscr{Z} \in \mathrm{Res}(s)$ with $z = \mathrm{corr}_{\mathscr{Z}}^{-1}(s)$ such that for each $i \in I$ we have $\Pr(\mathscr{C}^{\mathrm{t}}_{\max}(z, \Phi_i)) = r_i$.

We let $\mathbb{L}(s)$ denote the set of formulae satisfied by process $s \in \mathbf{S}$, namely $\mathbb{L}(s) = \{\Psi \in \mathbb{L}^{\mathrm{d}} \mid s \models \Psi\}$.

***Example* 4.** Consider process t in Fig. 3. It is easy to verify that $t \models 0.5\langle a \rangle \langle c \rangle \top \oplus 0.5 \langle a \rangle \langle b \rangle \top$. In fact, if we consider the resolution $\mathscr{Z}_t \in \mathrm{Res}(t)$ in the same figure, we have that the computation $c_1 = z_t \xrightarrow{a}\!\!\!\!\rightarrow z_{t_1} \xrightarrow{c}\!\!\!\!\rightarrow \mathrm{nil}$ is compatible with the trace formula $\langle a \rangle \langle c \rangle \top$ and that the computation $c_2 = z_t \xrightarrow{a}\!\!\!\!\rightarrow z_{t_2} \xrightarrow{b}\!\!\!\!\rightarrow \mathrm{nil}$ is compatible with the trace formula $\langle a \rangle \langle b \rangle \top$. Moreover, we have $\Pr(\mathscr{C}^{\mathrm{t}}_{\max}(z_t, \langle a \rangle \langle c \rangle \top)) = 0.5$ and $\Pr(\mathscr{C}^{\mathrm{t}}_{\max}(z_t, \langle a \rangle \langle b \rangle \top)) = 0.5$.

The modal logic $\mathbb{L}_{\mathrm{w}}$ differs from $\mathbb{L}$ solely in the labels of the diamonds in $\mathbb{L}^{\mathrm{t}}_{\mathrm{w}}$ which range over $\mathscr{A}_\tau$ in place of $\mathscr{A}$. Hence, syntax and semantics of $\mathbb{L}_{\mathrm{w}}$ directly follow from Definition 18 and Defs. 20-21, resp.

We let $\mathbb{L}_{\mathrm{w}}(s)$ denote the set of formlae satisfied by process $s \in \mathbf{S}$, namely $\mathbb{L}_{\mathrm{w}}(s) = \{\Psi \in \mathbb{L}^{\mathrm{d}}_{\mathrm{w}} \mid s \models \Psi\}$.

We introduce the $\mathbb{L}_{\mathrm{w}}$-*equivalence* which extends the equivalence of traces $\equiv_{\mathrm{w}}$ to trace formulae.

Definition 22 ($\mathbb{L}_{\mathrm{w}}$-equivalence of formulae)**.** The relation of $\mathbb{L}_{\mathrm{w}}$-*equivalence of trace formulae* $\equiv_{\mathrm{w}} \subseteq \mathbb{L}^{\mathrm{t}}_{\mathrm{w}} \times \mathbb{L}^{\mathrm{t}}_{\mathrm{w}}$ is the smallest equivalence relation satisfying (i) $\top \equiv_{\mathrm{w}} \top$ and (ii) $\langle \mathfrak{a}_1 \rangle \Phi_1 \equiv_{\mathrm{w}} \langle \mathfrak{a}_2 \rangle \Phi_2$ iff

- either $\mathfrak{a}_1 = \tau$ and $\Phi_1 \equiv_{\mathrm{w}} \langle \mathfrak{a}_2 \rangle \Phi_2$,
- or $\mathfrak{a}_2 = \tau$ and $\langle \mathfrak{a}_1 \rangle \Phi_1 \equiv_{\mathrm{w}} \Phi_2$
- or $\mathfrak{a}_1 = \mathfrak{a}_2$ and $\Phi_1 \equiv_{\mathrm{w}} \Phi_2$.

Then, the relation of $\mathbb{L}_{\mathrm{w}}$-*equivalence of trace distribution formulae* $\equiv^{\dagger}_{\mathrm{w}} \subseteq \mathbb{L}^{\mathrm{d}}_{\mathrm{w}} \times \mathbb{L}^{\mathrm{d}}_{\mathrm{w}}$ is obtained by lifting $\equiv_{\mathrm{w}}$ to a relation on probability distributions over trace formulae.

Remark 2. Clearly we have $\mathbb{L}_{\mathrm{w}/\equiv_{\mathrm{w}}} = \mathbb{L}$, namely the notion of $\equiv_{\mathrm{w}}$ coincides with the equality of formulae when restricted to $(\mathbb{L}^{\mathrm{d}} \times \mathbb{L}^{\mathrm{d}}) \cup (\mathbb{L}^{\mathrm{t}} \times \mathbb{L}^{\mathrm{t}})$. Given any $\Psi_1, \Psi_2 \in \mathbb{L}^{\mathrm{d}}$, we say that $\Psi_1 = \Psi_2$ if they express the same probability distribution over trace formulae.

Notice that we are using the same symbol $\equiv_{\mathrm{w}}$ to denote both the equivalence of traces and $\mathbb{L}_{\mathrm{w}}$-equivalence. The meaning will always be clear from the context.

5 Logical characterization of relations

In this section we present the characterization of strong (resp. weak) trace equivalence by means of $\mathbb{L}$ (resp. $\mathbb{L}_{\mathrm{w}}$) (Theorem 3 and Theorem 5). Following [8], we introduce the notion of *mimicking formula* of a resolution as a formula expressing the trace distribution for that resolution. Mimicking formulae characterize the (weak) trace distribution equivalence of resolutions: two resolutions are (weak) trace distribution equivalent iff their mimicking formulae are equal (resp. $\mathbb{L}_{\mathrm{w}}$-equivalent) (Theorem 2 and Theorem 4).

The *mimicking formula* of a resolution $\mathscr{Z} \in \mathrm{Res}(\mathbf{S})$ is defined as a trace distribution formula assigning a positive weight only to the maximal traces of $\mathscr{Z}$. Hence, we need to identify each maximal trace of $\mathscr{Z}$ with a proper trace formula. This is achieved through the notion of *tracing formula* of a trace.

Definition 23 (Tracing formula)**.** Given any trace $\alpha \in \mathscr{A}^\star$ we define the *tracing formula* of α, notation $\Phi_\alpha \in \mathbb{L}^{\mathrm{t}}$, inductively on the structure of α as follows:

$$\Phi_\alpha = \begin{cases} \top & \text{if } \alpha = \varepsilon \\ \langle a \rangle \Phi_{\alpha'} & \text{if } \alpha = a\alpha', \alpha' \in \mathscr{A}^\star. \end{cases}$$

Lemma 6. *Let $s \in \mathbf{S}$ and $\alpha \in \mathscr{A}^\star$. For each $c \in \mathscr{C}(s)$ we have $\mathrm{Tr}(c) = \alpha$ iff $c \models \Phi_\alpha$ and $|c| = \mathrm{dpt}(\Phi_\alpha)$.*

Proof. ($\Rightarrow$) Assume first that $\mathrm{Tr}(c) = \alpha$. We aim to show that this implies that $|c| = \mathrm{dpt}(\Phi_\alpha)$ and $c \models \Phi_\alpha$. To this aim we proceed by induction over $|c|$.

- Base case $|c| = 0$, namely c is the empty computation. Since $\alpha = \mathrm{Tr}(c)$, this gives that $\alpha = \varepsilon$ and therefore, by Def. 23, $\Phi_\varepsilon = \top$. Then from Def. 19 we gather $\mathrm{dpt}(\Phi_\alpha) = 0 = |c|$ and by Def. 20 we are guaranteed that $c \models \Phi_\varepsilon$.
- Inductive step $|c| > 0$. Assume wlog that $c = s \xrightarrow{a} c'$. In particular this implies that $|c'| < |c|$. Therefore, from $\alpha = \mathrm{Tr}(c)$ we get that α must be of the form $\alpha = a\alpha'$ for $\alpha' = \mathrm{Tr}(c')$. By Def. 23, $\alpha = a\alpha'$ implies $\Phi_\alpha = \langle a \rangle \Phi_{\alpha'}$. From $\alpha' = \mathrm{Tr}(c')$ and the inductive hypothesis over $|c'|$ we get that $\mathrm{dpt}(\Phi_{\alpha'}) = |c'|$ and $c' \models \Phi_{\alpha'}$. This, taken together with $c = s \xrightarrow{a} c'$ gives $c \models \Phi_\alpha$. Moreover, we have
$$\mathrm{dpt}(\Phi_\alpha) = \mathrm{dpt}(\Phi_{\alpha'}) + 1 = |c'| + 1 = |c|$$
thus concluding the proof.

($\Leftarrow$) Assume now that $|c| = \mathrm{dpt}(\Phi_\alpha)$ and $c \models \Phi_\alpha$. We aim to show that this implies that $\mathrm{Tr}(c) = \alpha$, namely that c is compatible with α. From $c \models \Phi_\alpha$ and the definition of tracing formula (Definition 23) we gather that the sequence of the labels of the first $\mathrm{dpt}(\Phi_\alpha)$ execution steps of c matches α. Moreover, $|c| = \mathrm{dpt}(\Phi_\alpha)$ implies that those steps are actually the only execution steps for c. Therefore we can immediately conclude that $\mathrm{Tr}(c) = \alpha$. □

We remark that a computation c is compatible with Φ_α iff c and α satisfy previous Lemma 6.

Definition 24 (Mimicking formula)**.** Consider any resolution $\mathscr{Z} \in \mathrm{Res}(\mathbf{S})$ with initial state z. We define the *mimicking formula* of $\mathscr{Z}$, notation $\Psi_{\mathscr{Z}}$, as

$$\Psi_{\mathscr{Z}} = \bigoplus_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))} \Pr(\mathscr{C}_{\max}(z, \alpha)) \Phi_\alpha$$

where, for each $\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))$, the formula Φ_α is the tracing formula of α.

Lemma 7. *For any resolution $\mathscr{Z} \in \mathrm{Res}(\mathbf{S})$, the mimicking formula of $\mathscr{Z}$ is a well defined trace distribution formula.*

Proof. By definition of mimicking formula (Definition 24) we have

$$\Psi_{\mathscr{Z}} = \bigoplus_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))} \Pr(\mathscr{C}_{\max}(z, \alpha)) \Phi_\alpha$$

where for each $\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))$ the formula Φ_α is the tracing formula of trace α.

Hence, to prove that $\Psi_{\mathscr{Z}}$ is a well defined trace distribution formula we simply need to show that

$$\sum_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))} \Pr(\mathscr{C}_{\max}(z, \alpha)) = 1$$

which follows by Lemma 2 by noticing that $\sum_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))} \Pr(\mathscr{C}_{\max}(z, \alpha)) = \sum_{\alpha \in \mathscr{A}^\star} \Pr(\mathscr{C}_{\max}(z, \alpha))$. □

***Example* 5.** Consider the resolutions $\mathscr{Z}_s \in \text{Res}(s)$ and $\mathscr{Z}_t \in \text{Res}(t)$ for processes s and t, resp., in Fig. 3. The mimicking formulae for them are, resp.

$$\Psi_{\mathscr{Z}_s} = 0.5\langle a\rangle\langle c\rangle\top \oplus 0.5\langle a\rangle\top$$
$$\Psi_{\mathscr{Z}_t} = 0.5\langle a\rangle\langle c\rangle\top \oplus 0.5\langle a\rangle\langle b\rangle\top.$$

The following results give us a first insight on the characterizing power of mimicking formulae: given $s \in \mathbf{S}$, the set of the mimicking formulae of its resolutions constitutes the set of formulae satisfied by s.

Proposition 7. *Let $s \in \mathbf{S}$. For each $\mathscr{Z} \in \text{Res}(s)$ it holds that $s \models \Psi_{\mathscr{Z}}$.*

Proof. Let $\mathscr{Z} \in \text{Res}(s)$, with $z = \text{corr}_{\mathscr{Z}}^{-1}(s)$. Hence, by definition of mimicking formula (Definition 24) we have that

$$\Psi_{\mathscr{Z}} = \bigoplus_{\alpha \in \text{Tr}(\mathscr{C}_{\max}(z))} \Pr(\mathscr{C}_{\max}(z,\alpha))\Phi_\alpha$$

where, for each $\alpha \in \text{Tr}(\mathscr{C}_{\max}(z))$ we have that Φ_α is the tracing formula of α. We need to show that $s \models \Psi_{\mathscr{Z}}$, namely we need to exhibit a resolution $\bar{\mathscr{Z}} \in \text{Res}(s)$, with $\bar{z} = \text{corr}_{\bar{\mathscr{Z}}}^{-1}(s)$, s.t. for each $\alpha \in \text{Tr}(\mathscr{C}_{\max}(z))$ we have that $\Pr(\mathscr{C}^{\text{t}}(\bar{z},\Phi_\alpha)) = \Pr(\mathscr{C}_{\max}(z,\alpha))$. We aim to show that $\mathscr{Z}$ is such a resolution, namely that for each $\alpha \in \text{Tr}(\mathscr{C}_{\max}(z))$ we have

$$\Pr(\mathscr{C}^{\text{t}}_{\max}(z,\Phi_\alpha)) = \Pr(\mathscr{C}_{\max}(z,\alpha)).$$

Let $\alpha \in \text{Tr}(\mathscr{C}_{\max}(z))$. By definition we have

$$\begin{aligned}
\mathscr{C}^{\text{t}}_{\max}(z,\Phi_\alpha) &= \{c \in \mathscr{C}_{\max}(z) \mid c \models \Phi_\alpha \wedge |c| = \text{dpt}(\Phi_\alpha)\} \\
&= \{c \in \mathscr{C}_{\max}(z) \mid \text{Tr}(c) = \alpha\} && \text{(by Lemma 6)} \\
&= \mathscr{C}_{\max}(z,\alpha) && (\alpha \in \text{Tr}(\mathscr{C}_{\max}(z))).
\end{aligned}$$

Thus, we can conclude that

$$\Pr(\mathscr{C}^{\text{t}}_{\max}(z,\Phi_\alpha)) = \sum_{c \in \mathscr{C}^{\text{t}}_{\max}(z,\Phi_\alpha)} \Pr(c) = \sum_{c \in \mathscr{C}_{\max}(z,\alpha)} \Pr(c) = \Pr(\mathscr{C}_{\max}(z,\alpha)).$$

□

Theorem 1. *Let $s \in \mathbf{S}$. We have that $\mathbb{L}(s) = \{1\top\} \cup \{\Psi_{\mathscr{Z}} \mid \mathscr{Z} \in \text{Res}(s)\}$.*

Proof. From Proposition 7 and the definition of the relation $\models$ (Definition 21) we can immediately infer that $\{\Psi_{\mathscr{Z}} \mid \mathscr{Z} \in \text{Res}(s)\} \subseteq \mathbb{L}(s)$. Moreover $1\top \in \mathbb{L}(s)$ is immediate. To conclude the proof we need to show that also the opposite inclusion holds, namely that $\mathbb{L}(s) \setminus \{1\top\} \subseteq \{\Psi_{\mathscr{Z}} \mid \mathscr{Z} \in \text{Res}(s)\}$. To this aim, consider any $\Psi = \bigoplus_{i \in I} r_i\Phi_i$ and assume that $\Psi \in \mathbb{L}(s)$. We have to show that Ψ is the mimicking formula of some resolution for s. Since $s \models \Psi$, from Definition 21 we can infer that there is at least one resolution $\mathscr{Z} \in \text{Res}(s)$ with $z = \text{corr}_{\mathscr{Z}}^{-1}(s)$ s.t. for each $i \in I$ we have $\Pr(\mathscr{C}^{\text{t}}_{\max}(z,\Phi_i)) = r_i$. We aim to prove that among the resolutions ensuring that $s \models \Psi$, there is a particular resolution $\mathscr{Z} \in \text{Res}(s)$ s.t.

$$\Psi = \Psi_{\mathscr{Z}}. \tag{2}$$

First of all we recall that by definition of trace distribution formula (Definition 18), for each $i \in I$ we have $r_i > 0$ and moreover $\sum_{i \in I} r_i = 1$. By definition of $\mathscr{C}^{\text{t}}$, we have that $c \in \mathscr{C}^{\text{t}}_{\max}(z,\Phi_i)$ iff $c \models \Phi_i$ and $|c| =$

$\mathrm{dpt}(\Phi_i)$, which by Lemma 6 implies that $\Phi_i = \Phi_{\mathrm{Tr}(c)}$. Hence, let us consider the resolution $\mathscr{Z} \in \mathrm{Res}(s)$ s.t. for each $i \in I$ we have $\mathscr{C}^{\mathrm{t}}_{\max}(z, \Phi_i) \subseteq \mathscr{C}_{\max}(z)$, namely the resolution s.t. the computations compatible with the trace formulae Φ_i are all maximal. Notice that the existence of such a resolution is guaranteed by $s \models \Psi$. Since for each $c \in \mathscr{C}^{\mathrm{t}}_{\max}(z, \Phi_i)$ we have $c \in \mathscr{C}_{\max}(z)$, we can infer that $\mathrm{Tr}(c) \in \mathrm{Tr}(\mathscr{C}_{\max}(z))$, namely $\Phi_i = \Phi_\alpha$ for some $\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))$. This gives that whenever $\Phi_i = \Phi_\alpha$, for some $\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))$, then we can prove (as done in the proof of Proposition 7) that

$$\mathrm{Pr}(\mathscr{C}^{\mathrm{t}}_{\max}(z, \Phi_i)) = \mathrm{Pr}(\mathscr{C}_{\max}(z, \alpha)). \tag{3}$$

Furthermore, we have obtained that $\{\Phi_i \mid i \in I\} \subseteq \{\Phi_\alpha \mid \alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))\}$.

To prove Equation (2) we need to show that also the opposite inclusion holds. Assume by contradiction that there is at least one $\beta \in \mathrm{Tr}(\mathscr{C}_{\max}(z))$ s.t. there is no $i \in I$ with $\Phi_i = \Phi_\beta$. Then we would have

$$\begin{array}{rll}
1 = & \sum_{i \in I} r_i & \\
= & \sum_{i \in I} \mathrm{Pr}(\mathscr{C}^{\mathrm{t}}_{\max}(z, \Phi_i)) & \\
\leq & \sum_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z)) \setminus \{\beta\}} \mathrm{Pr}(\mathscr{C}_{\max}(z, \alpha)) & \text{(by Equation (3))} \\
< & \sum_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))} \mathrm{Pr}(\mathscr{C}_{\max}(z, \alpha)) & (\beta \in \mathrm{Tr}(\mathscr{C}_{\max}(z)) \text{ implies } \mathrm{Pr}(\mathscr{C}_{\max}(z, \beta)) > 0) \\
= & 1 & \text{(by Lemma 2)}
\end{array}$$

which is a contradiction. Hence we can conclude that $\{\Phi_i \mid i \in I\} = \{\Phi_\alpha \mid \alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))\}$ and thus, due to Equation (3), that Equation (2) holds. □

Remark 3. In Theorem 1, $1\top$ is not included in the set of mimicking formulae of resolutions merely for sake of presentation, as $1\top$ is the mimicking formula of the resolution for s in which no action is executed.

The following theorem states that two resolutions are trace distribution equivalent iff their mimicking formulae are the same.

Theorem 2. *Let* $s, t \in \mathbf{S}$ *and consider* $\mathscr{Z}_s \in \mathrm{Res}(s)$*, with* $z_s = \mathrm{corr}^{-1}_{\mathscr{Z}_s}(s)$*, and* $\mathscr{Z}_t \in \mathrm{Res}(t)$*, with* $z_t = \mathrm{corr}^{-1}_{\mathscr{Z}_t}(t)$*. Then* $\Psi_{\mathscr{Z}_s} = \Psi_{\mathscr{Z}_t}$ *iff for all* $\alpha \in \mathscr{A}^\star$ *it holds that* $\mathrm{Pr}(\mathscr{C}(z_s, \alpha)) = \mathrm{Pr}(\mathscr{C}(z_t, \alpha))$.

Proof. ($\Rightarrow$) Assume first that $\Psi_{\mathscr{Z}_s} = \Psi_{\mathscr{Z}_t}$. We aim to show that this implies $\mathrm{Pr}(\mathscr{C}(z_s, \alpha)) = \mathrm{Pr}(\mathscr{C}(z_t, \alpha))$ for all $\alpha \in \mathscr{A}^\star$. By definition of mimicking formula (Definition 24) we have

$$\Psi_{\mathscr{Z}_s} = \bigoplus_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z_s))} \mathrm{Pr}(\mathscr{C}_{\max}(z_s, \alpha)) \Phi_\alpha$$

where for each $\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z_s))$ the formula Φ_α is the tracing formula of α. Analogously

$$\Psi_{\mathscr{Z}_t} = \bigoplus_{\beta \in \mathrm{Tr}(\mathscr{C}_{\max}(z_t))} \mathrm{Pr}(\mathscr{C}_{\max}(z_t, \beta)) \Phi_\beta$$

where for each $\beta \in \mathrm{Tr}(\mathscr{C}_{\max}(z_t))$ the formula Φ_β is the tracing formula of β.

Then from the assumption $\Psi_{\mathscr{Z}_s} = \Psi_{\mathscr{Z}_t}$ we gather

1. $\mathrm{Tr}(\mathscr{C}_{\max}(z_s)) = \mathrm{Tr}(\mathscr{C}_{\max}(z_t))$;
2. from previous item 1 we have that $\mathrm{Pr}(\mathscr{C}_{\max}(z_s, \alpha)) = \mathrm{Pr}(\mathscr{C}_{\max}(z_t, \alpha))$ for each $\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z_s))$.

We notice that item 1 above implies the stronger relation

$$\mathrm{Tr}(\mathscr{C}(z_s)) = \mathrm{Tr}(\mathscr{C}(z_t)). \tag{4}$$

In fact each $\alpha \in \mathrm{Tr}(\mathscr{C}(z_s))$ is either a trace in $\mathrm{Tr}(\mathscr{C}_{\max}(z_s))$ or a proper prefix of a trace in that set. In both cases item 1 guarantees that each trace in $\mathrm{Tr}(\mathscr{C}(z_s))$ has a matching trace in $\mathrm{Tr}(\mathscr{C}(z_t))$ and viceversa.

Now, consider any $\alpha \in \mathscr{A}^\star$. We aim to show that $\Pr(\mathscr{C}(z_s,\alpha)) = \Pr(\mathscr{C}(z_t,\alpha))$. For simplicity of presentation, we can distinguish two cases.

- $\Pr(\mathscr{C}(z_s,\alpha)) = 0$. In this case we have that no computation from z_s is compatible with α, namely there is no computation from z_s for which the sequence of the labels of the execution steps matches α. More precisely, we have that $\alpha \notin \mathrm{Tr}(\mathscr{C}(z_s))$. Since by Equation (4) we have that $\mathrm{Tr}(\mathscr{C}(z_s)) = \mathrm{Tr}(\mathscr{C}(z_t))$ we can directly conclude that $\alpha \notin \mathrm{Tr}(\mathscr{C}(z_t))$, namely $\Pr(\mathscr{C}(z_t,\alpha)) = 0$.
- $\Pr(\mathscr{C}(z_s,\alpha)) > 0$. In this case we have that $\alpha \in \mathrm{Tr}(\mathscr{C}(z_s))$ and by Equation (4) we have that this implies that $\alpha \in \mathrm{Tr}(\mathscr{C}(z_t))$. Hence we are guaranteed that $\Pr(\mathscr{C}(z_t,\alpha)) > 0$. It remains to show that $\Pr(\mathscr{C}(z_s,\alpha)) = \Pr(\mathscr{C}(z_t,\alpha))$. We have

$$\begin{aligned}
\Pr(\mathscr{C}(z_s,\alpha)) &= \textstyle\sum_{c \in P_{\max}(z_s,\alpha)} \Pr(c) && \text{(by Lemma 3)}\\
&= \textstyle\sum_{\beta \in \mathrm{Tr}(P_{\max}(z_s,\alpha))} \Pr(\mathscr{C}_{\max}(z_s,\beta)) && \text{(by def. of } P_{\max})\\
&= \textstyle\sum_{\beta \in \mathrm{Tr}(P_{\max}(z_s,\alpha))} \Pr(\mathscr{C}_{\max}(z_t,\beta)) && (P_{\max}(z_s,\alpha) \subseteq \mathscr{C}_{\max}(z_s) \text{ and item 2})\\
&= \textstyle\sum_{\beta' \in \mathrm{Tr}(P_{\max}(z_t,\alpha))} \Pr(\mathscr{C}_{\max}(z_t,\beta')) && \text{(by Equation (4))}\\
&= \textstyle\sum_{c' \in P_{\max}(z_t,\alpha)} \Pr(c') && \text{(by def. of } P_{\max})\\
&= \Pr(\mathscr{C}(z_t,\alpha)) && \text{(by Lemma 3).}
\end{aligned}$$

($\Leftarrow$) Assume now that for all $\alpha \in \mathscr{A}^\star$ it holds that $\Pr(\mathscr{C}(z_s,\alpha)) = \Pr(\mathscr{C}(z_t,\alpha))$. We aim to show that this implies that $\Psi_{\mathscr{Z}_s} = \Psi_{\mathscr{Z}_t}$. By definition of mimicking formula (Definition 24) we have

$$\Psi_{\mathscr{Z}_s} = \bigoplus_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z_s))} \Pr(\mathscr{C}_{\max}(z_s,\alpha))\Phi_\alpha$$

$$\Psi_{\mathscr{Z}_t} = \bigoplus_{\beta \in \mathrm{Tr}(\mathscr{C}_{\max}(z_t))} \Pr(\mathscr{C}_{\max}(z_t,\beta))\Phi_\beta.$$

Therefore, to prove $\Psi_{\mathscr{Z}_s} = \Psi_{\mathscr{Z}_t}$ we need to show that

$$\mathrm{Tr}(\mathscr{C}_{\max}(z_s)) = \mathrm{Tr}(\mathscr{C}_{\max}(z_t)) \tag{5}$$

$$\Pr(\mathscr{C}_{\max}(z_s,\alpha)) = \Pr(\mathscr{C}_{\max}(z_t,\alpha)) \text{ for each } \alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z_s)). \tag{6}$$

First of all we notice that $\Pr(\mathscr{C}(z_s,\alpha)) = \Pr(\mathscr{C}(z_t,\alpha))$ for each $\alpha \in \mathscr{A}^\star$ implies that $\mathrm{Tr}(\mathscr{C}(z_s)) = \mathrm{Tr}(\mathscr{C}(z_t))$. This is due to the fact that by definition, given any $\alpha \in \mathscr{A}^\star$, $\Pr(\mathscr{C}(z_s,\alpha)) > 0$ iff there is at least one computation $c \in \mathscr{C}(z_s)$ s.t. $\alpha = \mathrm{Tr}(c)$. Since $\Pr(\mathscr{C}(z_s,\alpha)) > 0$ implies $\Pr(\mathscr{C}(z_t,\alpha)) > 0$ we can infer that for each $\alpha \in \mathrm{Tr}(\mathscr{C}(z_s))$ there is at least one computation $c' \in \mathrm{Tr}(\mathscr{C}(z_t))$ s.t. $\alpha = \mathrm{Tr}(c')$, namely $\mathrm{Tr}(\mathscr{C}(z_s)) \subseteq \mathrm{Tr}(\mathscr{C}(z_t))$. As the same reasoning can be applied symmetrically to each $\alpha \in \mathrm{Tr}(\mathscr{C}(z_t))$, we can conclude that

$$\mathrm{Tr}(\mathscr{C}(z_s)) = \mathrm{Tr}(\mathscr{C}(z_t)). \tag{7}$$

Next we aim to show that a similar result holds even if we restrict our attention to maximal computations, that is we aim to prove Equation (5).

Let $\alpha \in \text{Tr}(\mathscr{C}_{\max}(z_s))$. Notice that for this α we have $\mathscr{C}_{\max}(z_s,\alpha) \subseteq P_{\max}(z_s,\alpha)$. Then we have

$$\begin{aligned}\Pr(\mathscr{C}(z_s,\alpha)) = &\ \textstyle\sum_{c\in P_{\max}(z_s,\alpha)} \Pr(c) && \text{(by Lemma 3)}\\ = &\ \textstyle\sum_{c\in \mathscr{C}_{\max}(z_s,\alpha)} \Pr(c) + \sum_{c'\in P_{\max}(z_s,\alpha)\setminus \mathscr{C}_{\max}(z_s,\alpha)} \Pr(c').\end{aligned} \tag{8}$$

Moreover, by Lemma 3 it holds that $\Pr(\mathscr{C}(z_t,\alpha)) = \sum_{c''\in P_{\max}(z_t,\alpha)} \Pr(c'')$.

Therefore, from $\Pr(\mathscr{C}(z_s,\alpha)) = \Pr(\mathscr{C}(z_t,\alpha))$ we gather that

$$\sum_{c\in \mathscr{C}_{\max}(z_s,\alpha)} \Pr(c) + \sum_{c'\in P_{\max}(z_s,\alpha)\setminus \mathscr{C}_{\max}(z_s,\alpha)} \Pr(c') = \sum_{c''\in P_{\max}(z_t,\alpha)} \Pr(c''). \tag{9}$$

Assume by contradiction that $P_{\max}(z_t,\alpha) \cap \mathscr{C}_{\max}(z_t,\alpha) = \emptyset$, namely there is no maximal computation from z_t which is compatible with α. Then for each action $a \in \mathscr{A}$ consider the trace αa and define $\text{Add}_{z_s}(\alpha) = \{a \in \mathscr{A} \mid \alpha a \in \text{Tr}(\mathscr{C}(z_s))\}$. From Equation (7) we can directly infer that $\text{Add}_{z_s}(\alpha) = \text{Add}_{z_t}(\alpha)$. Moreover, since we are assuming that no maximal computation from z_t is compatible with α, we get

$$\bigcup_{a\in \text{Add}_{z_s}(\alpha)} P_{\max}(z_s,\alpha a) = P_{\max}(z_s,\alpha) \setminus \mathscr{C}_{\max}(z_s,\alpha) \tag{10}$$

$$\bigcup_{a\in \text{Add}_{z_t}(\alpha)} P_{\max}(z_t,\alpha a) = P_{\max}(z_t,\alpha) \tag{11}$$

where the unions are guaranteed to be disjoint (a single computation cannot be compatible with more than one trace αa). Furthermore, by Lemma 3 we have that for each $a \in \text{Add}_{z_s}(\alpha)$

$$\begin{aligned}\Pr(\mathscr{C}(z_s,\alpha a)) = \textstyle\sum_{c_1\in P_{\max}(z_s,\alpha a)} \Pr(c_1)\\ \Pr(\mathscr{C}(z_t,\alpha a)) = \textstyle\sum_{c_2\in P_{\max}(z_t,\alpha a)} \Pr(c_2)\end{aligned}$$

from which we get that for each $a \in \text{Add}_{z_s}(\alpha)$ it holds that

$$\sum_{c_1\in P_{\max}(z_s,\alpha a)} \Pr(c_1) = \sum_{c_2\in P_{\max}(z_t,\alpha a)} \Pr(c_2). \tag{12}$$

Therefore we have that

$$\begin{aligned} & \textstyle\sum_{c\in \mathscr{C}_{\max}(z_s,\alpha)} \Pr(c) + \sum_{c'\in P_{\max}(z_s,\alpha)\setminus \mathscr{C}_{\max}(z_s,\alpha)} \Pr(c') \\ = & \textstyle\sum_{c\in \mathscr{C}_{\max}(z_s,\alpha)} \Pr(c) + \sum_{c'\in \bigcup_{a\in \text{Add}_{z_s}(\alpha)} P_{\max}(z_s,\alpha a)} \Pr(c') && \text{(by Equation (10))}\\ = & \textstyle\sum_{c\in \mathscr{C}_{\max}(z_s,\alpha)} \Pr(c) + \sum_{a\in \text{Add}_{z_s}(\alpha)} \left(\sum_{c'\in P_{\max}(z_s,\alpha a)} \Pr(c')\right) && \text{(disjoint union)}\\ = & \textstyle\sum_{c\in \mathscr{C}_{\max}(z_s,\alpha)} \Pr(c) + \sum_{a\in \text{Add}_{z_s}(\alpha)} \left(\sum_{c''\in P_{\max}(z_t,\alpha a)} \Pr(c'')\right) && \text{(by Equation (12))}\\ = & \textstyle\sum_{c\in \mathscr{C}_{\max}(z_s,\alpha)} \Pr(c) + \sum_{c''\in \bigcup_{a\in \text{Add}_{z_t}(\alpha)} P_{\max}(z_t,\alpha a)} \Pr(c'') && (\text{Add}_{z_s}(\alpha) = \text{Add}_{z_t}(\alpha) \text{ and disjoint union})\\ = & \textstyle\sum_{c\in \mathscr{C}_{\max}(z_s,\alpha)} \Pr(c) + \sum_{c''\in P_{\max}(z_t,\alpha)} \Pr(c'') && \text{(by Equation (11)).}\end{aligned}$$

Thus we have obtained that

$$\sum_{c\in \mathscr{C}_{\max}(z_s,\alpha)} \Pr(c) + \sum_{c'\in P_{\max}(z_s,\alpha)\setminus \mathscr{C}_{\max}(z_s,\alpha)} \Pr(c') = \sum_{c\in \mathscr{C}_{\max}(z_s,\alpha)} \Pr(c) + \sum_{c''\in P_{\max}(z_t,\alpha)} \Pr(c'')$$

which, since by the choice of α we have that $\sum_{c\in \mathscr{C}_{\max}(z_s,\alpha)} \Pr(c) > 0$, is in contradiction with Equation (9). Therefore, we have obtained that whenever $\alpha \in \text{Tr}(\mathscr{C}_{\max}(z_s))$ then there is at least one maximal computation c from z_t s.t. $\alpha = \text{Tr}(c)$, that is $\text{Tr}(\mathscr{C}_{\max}(z_s)) \subseteq \text{Tr}(\mathscr{C}_{\max}(z_t))$. Since the same reasoning can be applied symmetrically to each $\alpha \in \text{Tr}(\mathscr{C}_{\max}(z_t))$ we gather that also $\text{Tr}(\mathscr{C}_{\max}(z_t)) \subseteq \text{Tr}(\mathscr{C}_{\max}(z_s))$ holds. The two inclusions give us Equation (5).

Finally, we aim to prove Equation (6). Let $\alpha \in \text{Tr}(\mathscr{C}_{\max}(z_s))$. We can distinguish two cases.

- $|\alpha| = \mathrm{dpt}(z_s)$. First of all we notice that from Equation (7) and the assumption $\Pr(\mathscr{C}(z_s,\beta)) = \Pr(\mathscr{C}(z_t,\beta))$ for each $\beta \in \mathscr{A}^\star$, we can infer that $|\alpha| = \mathrm{dpt}(z_t)$. Hence, we have

$$\Pr(\mathscr{C}_{\max}(z_s,\alpha)) = \Pr(\mathscr{C}(z_s,\alpha)) = \Pr(\mathscr{C}(z_t,\alpha)) = \Pr(\mathscr{C}_{\max}(z_t,\alpha)).$$

- $|\alpha| < \mathrm{dpt}(z_s)$. Then we have

$$\begin{array}{cl}
& \Pr(\mathscr{C}_{\max}(z_s,\alpha)) \\
= & \sum_{c\in\mathscr{C}_{\max}(z_s,\alpha)} \Pr(c) \\
= & \sum_{c'\in P_{\max}(z_s,\alpha)} \Pr(c') - \sum_{c''\in P_{\max}(z_s,\alpha)\setminus\mathscr{C}_{\max}(z_s,\alpha)} \Pr(c'') \\
= & \Pr(\mathscr{C}(z_s,\alpha)) - \sum_{c''\in P_{\max}(z_s,\alpha)\setminus\mathscr{C}_{\max}(z_s,\alpha)} \Pr(c'') \\
= & \Pr(\mathscr{C}(z_t,\alpha)) - \sum_{c''\in P_{\max}(z_s,\alpha)\setminus\mathscr{C}_{\max}(z_s,\alpha)} \Pr(c'') \\
= & \sum_{c'''\in P_{\max}(z_t,\alpha)} \Pr(c''') - \sum_{c''\in P_{\max}(z_s,\alpha)\setminus\mathscr{C}_{\max}(z_s,\alpha)} \Pr(c'') \\
= & \sum_{c_1\in\mathscr{C}_{\max}(z_t,\alpha)} \Pr(c_1) + \sum_{c_2\in P_{\max}(z_t,\alpha)\setminus\mathscr{C}_{\max}(z_t,\alpha)} \Pr(c_2) - \sum_{c''\in P_{\max}(z_s,\alpha)\setminus\mathscr{C}_{\max}(z_s,\alpha)} \Pr(c'') \\
= & \sum_{c_1\in\mathscr{C}_{\max}(z_t,\alpha)} \Pr(c_1) + \sum_{c_2\in\bigcup_{b\in\mathrm{Add}_{z_t}(\alpha)} P_{\max}(z_t,\alpha b)} \Pr(c_2) - \sum_{c''\in\bigcup_{b\in\mathrm{Add}_{z_s}(\alpha)} P_{\max}(z_s,\alpha b)} \Pr(c'') \\
= & \sum_{c_1\in\mathscr{C}_{\max}(z_t,\alpha)} \Pr(c_1) + \sum_{b\in\mathrm{Add}_{z_t}(\alpha)} \left(\sum_{c_2\in P_{\max}(z_t,\alpha b)} \Pr(c_2)\right) - \sum_{b\in\mathrm{Add}_{z_s}(\alpha)} \left(\sum_{c''\in P_{\max}(z_s,\alpha b)} \Pr(c'')\right) \\
= & \sum_{c_1\in\mathscr{C}_{\max}(z_t,\alpha)} \Pr(c_1) + \sum_{b\in\mathrm{Add}_{z_t}(\alpha)} \Pr(\mathscr{C}(z_t,\alpha b)) - \sum_{b\in\mathrm{Add}_{z_s}(\alpha)} \Pr(\mathscr{C}(z_s,\alpha b)) \\
= & \sum_{c_1\in\mathscr{C}_{\max}(z_t,\alpha)} \Pr(c_1) \\
= & \Pr(\mathscr{C}_{\max}(z_t,\alpha))
\end{array}$$

  where
  - the second and the sixth steps follow by Equation (8);
  - the third, fifth and ninth steps follow by Lemma 3;
  - the fourth step follows by the initial assumption which guarantees that $\Pr(\mathscr{C}(z_s,\alpha)) = \Pr(\mathscr{C}(z_t,\alpha))$;
  - the seventh step follows by Equation (10);
  - the tenth step follows by $\mathrm{Add}_{z_s}(\alpha) = \mathrm{Add}_{z_t}(\alpha)$ (given by Equation (7)) and the initial assumption which guarantees that for each $b \in \mathrm{Add}_{z_s}(\alpha)$, $\Pr(\mathscr{C}(z_s,\alpha b)) = \Pr(\mathscr{C}(z_t,\alpha b))$.

□

Then we can derive the characterization result for the strong case: two processes s,t are strong trace equivalent iff they satisfy the same formulae in $\mathbb{L}$.

Theorem 3. *For all $s,t \in \mathbf{S}$ we have that $s \approx_{\mathrm{st}} t$ iff $\mathbb{L}(s) = \mathbb{L}(t)$.*

Proof. ($\Rightarrow$) Assume first that $s \approx_{\mathrm{st}} t$. We aim to sow that this implies that $\mathbb{L}(s) = \mathbb{L}(t)$. By Definition 6 $s \approx_{\mathrm{st}} t$ implies that

(i) for each resolution $\mathscr{Z}_s \in \mathrm{Res}(s)$, with $z_s = \mathrm{corr}^{-1}_{\mathscr{Z}_s}(s)$, there is a resolution $\mathscr{Z}_t \in \mathrm{Res}(t)$, with $z_t = \mathrm{corr}^{-1}_{\mathscr{Z}_t}(t)$, s.t. for each $\alpha \in \mathscr{A}^\star$ we have $\Pr(\mathscr{C}(z_s,\alpha)) = \Pr(\mathscr{C}(z_t,\alpha))$;

(ii) for each resolution $\mathscr{Z}_t \in \mathrm{Res}(t)$, with $z_t = \mathrm{corr}^{-1}_{\mathscr{Z}_t}(t)$, there is a resolution $\mathscr{Z}_s \in \mathrm{Res}(s)$, with $z_s = \mathrm{corr}^{-1}_{\mathscr{Z}_s}(s)$, s.t. for each $\alpha \in \mathscr{A}^\star$ we have $\Pr(\mathscr{C}(z_s,\alpha)) = \Pr(\mathscr{C}(z_t,\alpha))$.

Consider any $\mathscr{Z}_s \in \mathrm{Res}(s)$, with $z_s = \mathrm{corr}^{-1}_{\mathscr{Z}_s}(s)$, and let $\mathscr{Z}_t \in \mathrm{Res}(t)$, with $z_t = \mathrm{corr}^{-1}_{\mathscr{Z}_t}(t)$, be any resolution of t satisfying item (i) above. By Theorem 2, $\Pr(\mathscr{C}(z_s,\alpha)) = \Pr(\mathscr{C}(z_t,\alpha))$ for all $\alpha \in \mathscr{A}^\star$ implies that $\Psi_{\mathscr{Z}_s} = \Psi_{\mathscr{Z}_t}$. More precisely, we have that

$$\text{for each } \mathscr{Z}_s \in \mathrm{Res}(s) \text{ there is } \mathscr{Z}_t \in \mathrm{Res}(t) \text{ s.t. } \Psi_{\mathscr{Z}_s} = \Psi_{\mathscr{Z}_t}. \tag{13}$$

Symmetrically, item (ii) above taken together with Theorem 2 gives that

$$\text{for each } \mathscr{Z}_t \in \mathrm{Res}(t) \text{ there is a } \mathscr{Z}_s \in \mathrm{Res}(s) \text{ s.t. } \Psi_{\mathscr{Z}_t} = \Psi_{\mathscr{Z}_s}. \tag{14}$$

Therefore, from Equations (13) and (14) we gather

$$\{\Psi_{\mathscr{Z}_s} \mid \mathscr{Z}_s \in \mathrm{Res}(s)\} = \{\Psi_{\mathscr{Z}_t} \mid \mathscr{Z}_t \in \mathrm{Res}(t)\}. \tag{15}$$

By Theorem 1 we have that $\mathbb{L}(s) = \{1\top\} \cup \{\Psi_{\mathscr{Z}_s} \mid \mathscr{Z}_s \in \mathrm{Res}(s)\}$ and similarly $\mathbb{L}(t) = \{1\top\} \cup \{\Psi_{\mathscr{Z}_t} \mid \mathscr{Z}_t \in \mathrm{Res}(t)\}$. Therefore, from Equation (15) we can conclude that $\mathbb{L}(s) = \mathbb{L}(t)$.

($\Leftarrow$) Assume now that $\mathbb{L}(s) = \mathbb{L}(t)$. We aim to show that this implies that $s \approx_{\mathrm{st}} t$. By Theorem 1 we have that $\mathbb{L}(s) = \{1\top\} \cup \{\Psi_{\mathscr{Z}_s} \mid \mathscr{Z}_s \in \mathrm{Res}(s)\}$ and analogously $\mathbb{L}(t) = \{1\top\} \cup \{\Psi_{\mathscr{Z}_t} \mid \mathscr{Z}_t \in \mathrm{Res}(t)\}$. Hence, from the assumption we can infer that $\{\Psi_{\mathscr{Z}_s} \mid \mathscr{Z}_s \in \mathrm{Res}(s)\} = \{\Psi_{\mathscr{Z}_t} \mid \mathscr{Z}_t \in \mathrm{Res}(t)\}$.

Clearly the equality between the two sets implies that

- for each $\mathscr{Z}_s \in \mathrm{Res}(s)$ there is a $\mathscr{Z}_t \in \mathrm{Res}(t)$ s.t. $\Psi_{\mathscr{Z}_s} = \Psi_{\mathscr{Z}_t}$ and
- for each $\mathscr{Z}_t \in \mathrm{Res}(t)$ there is a $\mathscr{Z}_s \in \mathrm{Res}(s)$ s.t. $\Psi_{\mathscr{Z}_t} = \Psi_{\mathscr{Z}_s}$.

By applying Theorem 2 to the two items above we obtain that

- for each resolution $\mathscr{Z}_s \in \mathrm{Res}(s)$, with $z_s = \mathrm{corr}_{\mathscr{Z}_s}^{-1}(s)$, there is a resolution $\mathscr{Z}_t \in \mathrm{Res}(t)$, with $z_t = \mathrm{corr}_{\mathscr{Z}_t}^{-1}(t)$, s.t. for each $\alpha \in \mathscr{A}^\star$ we have $\Pr(\mathscr{C}(z_s, \alpha)) = \Pr(\mathscr{C}(z_t, \alpha))$;
- for each resolution $\mathscr{Z}_t \in \mathrm{Res}(t)$, with $z_t = \mathrm{corr}_{\mathscr{Z}_t}^{-1}(t)$, there is a resolution $\mathscr{Z}_s \in \mathrm{Res}(s)$, with $z_s = \mathrm{corr}_{\mathscr{Z}_s}^{-1}(s)$, s.t. for each $\alpha \in \mathscr{A}^\star$ we have $\Pr(\mathscr{C}(z_s, \alpha)) = \Pr(\mathscr{C}(z_t, \alpha))$;

from which we can conclude that $s \approx_{\mathrm{st}} t$. □

The notions of *tracing formula* and *mimicking formula* and the related results Lemma 6, Lemma 7, Proposition 7 and Theorem 1 can be easily extended to the weak case by extending the set of traces $\mathscr{A}^\star$ to the set $\mathscr{A}_\tau^\star$.

The following theorem gives the characterization of weak trace distribution equivalence: two resolutions are weak trace distribution equivalent iff their mimicking formulae are $\mathbb{L}_{\mathrm{w}}$-equivalent.

To simplify the upcoming proofs, we introduce an alternative version of the *weak mimicking formula*, which captures the weak trace distribution (see Definition 16) of resolutions.

Definition 25. Consider any resolution $\mathscr{Z} \in \mathrm{Res}(\mathbf{S})$ with initial state z. We define the *weak mimicking formula* of $\mathscr{Z}$ as the trace distribution formula $\Psi^{\mathrm{w}}_{\mathscr{Z}}$ given by

$$\Psi^{\mathrm{w}}_{\mathscr{Z}} = \bigoplus_{\alpha \in \mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z))} \Pr(\mathscr{C}^{\mathrm{w}}_{\max}(z, \alpha)) \Phi_\alpha$$

where, for each $\alpha \in \mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z))$, the formula Φ_α is the tracing formula of α.

Notice that from the definitions of $\mathscr{C}^{\mathrm{w}}_{\max}(\_,\_)$ and $\mathrm{Tr}_{\mathrm{w}}(\_)$ we can infer that $\Psi^{\mathrm{w}}_{\_}$ represents a trace distribution formula over the quotient space of $\mathbb{L}_{\mathrm{w}}$ wrt. $\equiv_{\mathrm{w}}$, that is $\Psi^{\mathrm{w}}_{\_} \in \mathbb{L}^{\mathrm{d}}$.

Lemma 8. *For each $\mathscr{Z} \in \mathrm{Res}(\mathbf{S})$ it holds that $\Psi_{\mathscr{Z}} \equiv^{\dagger}_{\mathrm{w}} \Psi^{\mathrm{w}}_{\mathscr{Z}}$.*

Proof. Consider $\mathscr{Z} \in \mathrm{Res}(\mathbf{S})$ with initial state z. First of all we recall that by definition of mimicking formula (Definition 24) we have

$$\Psi_{\mathscr{Z}} = \bigoplus_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))} \Pr(\mathscr{C}_{\max}(z, \alpha)) \Phi_\alpha$$

where for each $\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))$, the formula Φ_α is the tracing formula of α. By definition of weak mimicking formula (Definition 25) we have

$$\Psi^{\mathrm{w}}_{\mathscr{Z}} = \bigoplus_{\beta \in \mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z))} \mathrm{Pr}(\mathscr{C}^{\mathrm{w}}_{\max}(z,\beta))\Phi_\beta$$

where for each $\beta \in \mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z))$, the formula Φ_β is the tracing formula of β. Moreover, we have that for each $\beta \in \mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z))$

$$\begin{aligned}
\mathrm{Pr}(\mathscr{C}^{\mathrm{w}}_{\max}(z,\beta)) = {} & \textstyle\sum_{c \in \mathscr{C}^{\mathrm{w}}_{\max}(z,\beta)} \mathrm{Pr}(c) \\
= {} & \textstyle\sum_{c \in \mathscr{C}_{\max}(z) \text{ s.t. } \mathrm{Tr}(c) \equiv_{\mathrm{w}} \beta} \mathrm{Pr}(c) \\
= {} & \textstyle\sum_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z)) \text{ s.t. } \alpha \equiv_{\mathrm{w}} \beta} \mathrm{Pr}(\mathscr{C}_{\max}(z,\alpha)).
\end{aligned}$$

Furthermore, by definition of tracing formula (Definition 23) and of $\equiv_{\mathrm{w}}$ (Definition 7), it is immediate that $\alpha \equiv_{\mathrm{w}} \beta$ iff $\Phi_\alpha \equiv_{\mathrm{w}} \Phi_\beta$, for each $\alpha, \beta \in \mathscr{A}^\star$. For simplicity, we denote by α_β each $\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))$ s.t. $\alpha \equiv_{\mathrm{w}} \beta$ for some $\beta \in \mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z))$. Notice that by construction of $\mathrm{Tr}_{\mathrm{w}}(\_)$, no trace $\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))$ can be equivalent to more than one $\beta \in \mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z))$. Therefore, we have obtained that

$$\begin{aligned}
\Psi^{\mathrm{w}}_{\mathscr{Z}} = {} & \textstyle\bigoplus_{\beta \in \mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z))} \mathrm{Pr}(\mathscr{C}^{\mathrm{w}}_{\max}(z,\beta))\Phi_\beta \\
\equiv^{\dagger}_{\mathrm{w}} {} & \textstyle\bigoplus_{\substack{\beta \in \mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z)) \\ \alpha_\beta \in \mathrm{Tr}(\mathscr{C}_{\max}(z))}} \mathrm{Pr}(\mathscr{C}_{\max}(z,\alpha_\beta))\Phi_{\alpha_\beta} \\
\equiv^{\dagger}_{\mathrm{w}} {} & \textstyle\bigoplus_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z))} \mathrm{Pr}(\mathscr{C}_{\max}(z,\alpha))\Phi_\alpha \\
= {} & \Psi_{\mathscr{Z}}.
\end{aligned}$$

□

Theorem 4. *Let $s,t \in \mathbf{S}$ and consider $\mathscr{Z}_s \in \mathrm{Res}(s)$, with $z_s = \mathrm{corr}^{-1}_{\mathscr{Z}_s}(s)$, and $\mathscr{Z}_t \in \mathrm{Res}(t)$, with $z_t = \mathrm{corr}^{-1}_{\mathscr{Z}_t}(t)$. Then $\Psi_{\mathscr{Z}_s} \equiv^{\dagger}_{\mathrm{w}} \Psi_{\mathscr{Z}_t}$ iff for all $\alpha \in \mathscr{A}^\star$ it holds that $\mathrm{Pr}(\mathscr{C}^{\mathrm{w}}(z_s,\alpha)) = \mathrm{Pr}(\mathscr{C}^{\mathrm{w}}(z_t,\alpha))$.*

Proof. ($\Rightarrow$) Assume first that $\Psi_{\mathscr{Z}_s} \equiv^{\dagger}_{\mathrm{w}} \Psi_{\mathscr{Z}_t}$. We aim to show that this implies $\mathrm{Pr}(\mathscr{C}^{\mathrm{w}}(z_s,\alpha)) = \mathrm{Pr}(\mathscr{C}^{\mathrm{w}}(z_t,\alpha))$ for all $\alpha \in \mathscr{A}^\star$. By Lemma 8 we have that

$$\Psi_{\mathscr{Z}_s} \equiv^{\dagger}_{\mathrm{w}} \Psi^{\mathrm{w}}_{\mathscr{Z}_s} \quad \text{and} \quad \Psi_{\mathscr{Z}_t} \equiv^{\dagger}_{\mathrm{w}} \Psi^{\mathrm{w}}_{\mathscr{Z}_t}.$$

Thus, $\Psi_{\mathscr{Z}_s} \equiv^{\dagger}_{\mathrm{w}} \Psi_{\mathscr{Z}_t}$ implies $\Psi^{\mathrm{w}}_{\mathscr{Z}_s} \equiv^{\dagger}_{\mathrm{w}} \Psi^{\mathrm{w}}_{\mathscr{Z}_t}$. Hence the prove the proof obligation, it is enough to prove that

$$\Psi^{\mathrm{w}}_{\mathscr{Z}_s} \equiv^{\dagger}_{\mathrm{w}} \Psi^{\mathrm{w}}_{\mathscr{Z}_t} \text{ implies } \mathrm{Pr}(\mathscr{C}^{\mathrm{w}}(z_s,\alpha)) = \mathrm{Pr}(\mathscr{C}^{\mathrm{w}}(z_t,\alpha)) \text{ for each } \alpha \in \mathscr{A}^\star. \tag{16}$$

From $\Psi^{\mathrm{w}}_{\mathscr{Z}_s} \equiv^{\dagger}_{\mathrm{w}} \Psi^{\mathrm{w}}_{\mathscr{Z}_t}$ we get that

$$\Psi^{\mathrm{w}}_{\mathscr{Z}_t} = \bigoplus_{\substack{\alpha \in \mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z_s)) \\ \beta_\alpha \in \mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z_t)) \cap [\alpha]_{\mathrm{w}}}} \mathrm{Pr}(\mathscr{C}^{\mathrm{w}}_{\max}(z_t,\beta_\alpha))\Phi_{\beta_\alpha}$$

with $\sum_{\beta_\alpha \in \mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z_t)) \cap [\alpha]_{\mathrm{w}}} \mathrm{Pr}(\mathscr{C}^{\mathrm{w}}_{\max}(z_t,\beta_\alpha)) = \mathrm{Pr}(\mathscr{C}^{\mathrm{w}}_{\max}(z_s,\alpha))$ and $\Phi_{\beta_\alpha} \equiv_{\mathrm{w}} \Phi_\alpha$ for each $\beta_\alpha \in \mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z_t)) \cap [\alpha]_{\mathrm{w}}, \alpha \in \mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z_s))$.

We notice that by definition the elements of $\mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z_t))$ represent distinct equivalence classes with respect to $\equiv_{\mathrm{w}}$. Thus we are guaranteed that for each $\alpha \in \mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z_t)) \cap [\alpha]_{\mathrm{w}}$ contains a single trace β_α. Therefore, in this particular case, $\Psi^{\mathrm{w}}_{\mathscr{Z}_s} \equiv^{\dagger}_{\mathrm{w}} \Psi^{\mathrm{w}}_{\mathscr{Z}_t}$ is equivalent to say that $\Psi^{\mathrm{w}}_{\mathscr{Z}_s} = \Psi^{\mathrm{w}}_{\mathscr{Z}_t}$. Moreover, since the representative of the equivalence classes wrt $\equiv_{\mathrm{w}}$ can always be chosen in $\mathscr{A}^\star$, we can always construct the sets $\mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z_s))$ and $\mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z_t))$ in such a way that $\mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z_s)) \cap \mathscr{A}^\star = \mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z_s))$ and $\mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z_t)) \cap \mathscr{A}^\star = \mathrm{Tr}_{\mathrm{w}}(\mathscr{C}_{\max}(z_t))$. Hence, the same argumentations presented in the first part of the proof of Theorem 2 allow us to prove the proof obligation Equation (16).

($\Leftarrow$) Assume now that for all $\alpha \in \mathscr{A}^\star$ it holds that $\Pr(\mathscr{C}^{\mathrm{w}}(z_s,\alpha)) = \Pr(\mathscr{C}^{\mathrm{w}}(z_t,\alpha))$. We aim to show that this implies that $\Psi_{\mathscr{Z}_s} \equiv_{\mathrm{w}}^{\dagger} \Psi_{\mathscr{Z}_t}$. To this aim we show that the assumption $\Pr(\mathscr{C}^{\mathrm{w}}(z_s,\alpha)) = \Pr(\mathscr{C}^{\mathrm{w}}(z_t,\alpha))$ for all $\alpha \in \mathscr{A}^\star$ implies $\Psi^{\mathrm{w}}_{\mathscr{Z}_s} = \Psi^{\mathrm{w}}_{\mathscr{Z}_t}$. This follows from the same argumentations presented in the second part of the proof of Theorem 2. Then, since from Lemma 8 we have $\Psi_{\mathscr{Z}_s} \equiv_{\mathrm{w}}^{\dagger} \Psi^{\mathrm{w}}_{\mathscr{Z}_s}$ and $\Psi_{\mathscr{Z}_t} \equiv_{\mathrm{w}}^{\dagger} \Psi^{\mathrm{w}}_{\mathscr{Z}_t}$, we can conclude that $\Psi_{\mathscr{Z}_s} \equiv_{\mathrm{w}}^{\dagger} \Psi_{\mathscr{Z}_t}$ as required. □

Then we can derive the characterization result for the weak case: two processes s,t are weak trace equivalent iff they satisfy equivalent formulae in $\mathbb{L}_{\mathrm{w}}$.

Theorem 5. *For all $s,t \in \mathbf{S}$ we have that $s \approx_{\mathrm{wt}} t$ iff $\mathbb{L}_{\mathrm{w}}(s) \equiv_{\mathrm{w}}^{\dagger} \mathbb{L}_{\mathrm{w}}(t)$.*

Proof. ($\Rightarrow$) Assume first that $s \approx_{\mathrm{wt}} t$. We aim to sow that this implies that $\mathbb{L}_{\mathrm{w}}(s) \equiv_{\mathrm{w}}^{\dagger} \mathbb{L}_{\mathrm{w}}(t)$. By Definition 8 $s \approx_{\mathrm{wt}} t$ implies that

(i) for each resolution $\mathscr{Z}_s \in \mathrm{Res}(s)$, with $z_s = \mathrm{corr}_{\mathscr{Z}_s}^{-1}(s)$, there is a resolution $\mathscr{Z}_t \in \mathrm{Res}(t)$, with $z_t = \mathrm{corr}_{\mathscr{Z}_t}^{-1}(t)$, s.t. for each $\alpha \in \mathscr{A}^\star$ we have $\Pr(\mathscr{C}^{\mathrm{w}}(z_s,\alpha)) = \Pr(\mathscr{C}^{\mathrm{w}}(z_t,\alpha))$;

(ii) for each resolution $\mathscr{Z}_t \in \mathrm{Res}(t)$, with $z_t = \mathrm{corr}_{\mathscr{Z}_t}^{-1}(t)$, there is a resolution $\mathscr{Z}_s \in \mathrm{Res}(s)$, with $z_s = \mathrm{corr}_{\mathscr{Z}_s}^{-1}(s)$, s.t. for each $\alpha \in \mathscr{A}^\star$ we have $\Pr(\mathscr{C}^{\mathrm{w}}(z_s,\alpha)) = \Pr(\mathscr{C}^{\mathrm{w}}(z_t,\alpha))$.

Consider any $\mathscr{Z}_s \in \mathrm{Res}(s)$, with $z_s = \mathrm{corr}_{\mathscr{Z}_s}^{-1}(s)$, and let $\mathscr{Z}_t \in \mathrm{Res}(t)$, with $z_t = \mathrm{corr}_{\mathscr{Z}_t}^{-1}(t)$, be any resolution of t satisfying item (i) above. By Theorem 4, $\Pr(\mathscr{C}^{\mathrm{w}}(z_s,\alpha)) = \Pr(\mathscr{C}^{\mathrm{w}}(z_t,\alpha))$ for all $\alpha \in \mathscr{A}^\star$ implies that $\Psi_{\mathscr{Z}_s} \equiv_{\mathrm{w}}^{\dagger} \Psi_{\mathscr{Z}_t}$. More precisely, we have that

$$\text{for each } \mathscr{Z}_s \in \mathrm{Res}(s) \text{ there is } \mathscr{Z}_t \in \mathrm{Res}(t) \text{ s.t. } \Psi_{\mathscr{Z}_s} \equiv_{\mathrm{w}}^{\dagger} \Psi_{\mathscr{Z}_t}. \tag{17}$$

Symmetrically, item (ii) above taken together with Theorem 4 gives that

$$\text{for each } \mathscr{Z}_t \in \mathrm{Res}(t) \text{ there is a } \mathscr{Z}_s \in \mathrm{Res}(s) \text{ s.t. } \Psi_{\mathscr{Z}_t} \equiv_{\mathrm{w}}^{\dagger} \Psi_{\mathscr{Z}_s}. \tag{18}$$

Therefore, from Equations (17) and (18) we gather

$$\{\Psi_{\mathscr{Z}_s} \mid \mathscr{Z}_s \in \mathrm{Res}(s)\} \equiv_{\mathrm{w}}^{\dagger} \{\Psi_{\mathscr{Z}_t} \mid \mathscr{Z}_t \in \mathrm{Res}(t)\}. \tag{19}$$

By Theorem 1 we have that $\mathbb{L}_{\mathrm{w}}(s) = \{1\top\} \cup \{\Psi_{\mathscr{Z}_s} \mid \mathscr{Z}_s \in \mathrm{Res}(s)\}$ and similarly $\mathbb{L}_{\mathrm{w}}(t) = \{1\top\} \cup \{\Psi_{\mathscr{Z}_t} \mid \mathscr{Z}_t \in \mathrm{Res}(t)\}$. Therefore, from Equation (19) we can conclude that $\mathbb{L}_{\mathrm{w}}(s) \equiv_{\mathrm{w}}^{\dagger} \mathbb{L}_{\mathrm{w}}(t)$.

($\Leftarrow$) Assume now that $\mathbb{L}_{\mathrm{w}}(s) \equiv_{\mathrm{w}}^{\dagger} \mathbb{L}_{\mathrm{w}}(t)$. We aim to show that this implies that $s \approx_{\mathrm{wt}} t$. By Theorem 1 we have that $\mathbb{L}_{\mathrm{w}}(s) = \{1\top\} \cup \{\Psi_{\mathscr{Z}_s} \mid \mathscr{Z}_s \in \mathrm{Res}(s)\}$ and analogously $\mathbb{L}_{\mathrm{w}}(t) = \{1\top\} \cup \{\Psi_{\mathscr{Z}_t} \mid \mathscr{Z}_t \in \mathrm{Res}(t)\}$. Hence, from the assumption we can infer that $\{\Psi_{\mathscr{Z}_s} \mid \mathscr{Z}_s \in \mathrm{Res}(s)\} \equiv_{\mathrm{w}}^{\dagger} \{\Psi_{\mathscr{Z}_t} \mid \mathscr{Z}_t \in \mathrm{Res}(t)\}$.

Clearly the equivalence between the two sets implies that

- for each $\mathscr{Z}_s \in \mathrm{Res}(s)$ there is a $\mathscr{Z}_t \in \mathrm{Res}(t)$ s.t. $\Psi_{\mathscr{Z}_s} \equiv_{\mathrm{w}}^{\dagger} \Psi_{\mathscr{Z}_t}$ and
- for each $\mathscr{Z}_t \in \mathrm{Res}(t)$ there is a $\mathscr{Z}_s \in \mathrm{Res}(s)$ s.t. $\Psi_{\mathscr{Z}_t} \equiv_{\mathrm{w}}^{\dagger} \Psi_{\mathscr{Z}_s}$.

By applying Theorem 4 to the two items above we obtain that

- for each resolution $\mathscr{Z}_s \in \mathrm{Res}(s)$, with $z_s = \mathrm{corr}_{\mathscr{Z}_s}^{-1}(s)$, there is a resolution $\mathscr{Z}_t \in \mathrm{Res}(t)$, with $z_t = \mathrm{corr}_{\mathscr{Z}_t}^{-1}(t)$, s.t. for each $\alpha \in \mathscr{A}^\star$ we have $\Pr(\mathscr{C}^{\mathrm{w}}(z_s,\alpha)) = \Pr(\mathscr{C}^{\mathrm{w}}(z_t,\alpha))$;
- for each resolution $\mathscr{Z}_t \in \mathrm{Res}(t)$, with $z_t = \mathrm{corr}_{\mathscr{Z}_t}^{-1}(t)$, there is a resolution $\mathscr{Z}_s \in \mathrm{Res}(s)$, with $z_s = \mathrm{corr}_{\mathscr{Z}_s}^{-1}(s)$, s.t. for each $\alpha \in \mathscr{A}^\star$ we have $\Pr(\mathscr{C}^{\mathrm{w}}(z_s,\alpha)) = \Pr(\mathscr{C}^{\mathrm{w}}(z_t,\alpha))$;

from which we can conclude that $s \approx_{\mathrm{wt}} t$. □

6 Logical characterization of trace metrics

In this section we present the logical characterization of strong and weak trace metric (resp. Theorem 8 and Theorem 11). We define a suitable distance on formulae in $\mathbb{L}$ (resp. $\mathbb{L}_w$) and we characterize the strong (resp. weak) trace metric between processes as the distance between the sets of formulae satisfied by them.

6.1 $\mathbb{L}$-characterization of strong trace metric

Firstly, we need to define a distance on trace formulae.

Definition 26 (Distance on $\mathbb{L}^t$)**.** The function $\mathscr{D}_{\mathbb{L}}^{t}\colon \mathbb{L}^t \times \mathbb{L}^t \to [0,1]$ is defined over $\mathbb{L}^t$ as follows:

$$\mathscr{D}_{\mathbb{L}}^{t}(\Phi_1,\Phi_2) = \begin{cases} 0 & \text{if } \Phi_1 = \Phi_2 \\ 1 & \text{otherwise.} \end{cases}$$

Proposition 8. *The function $\mathscr{D}_{\mathbb{L}}^{t}$ is a 1-bounded metric over $\mathbb{L}^t$.*

Proof. The thesis follows by noticing that $\mathscr{D}_{\mathbb{L}}^{t}$ is the discrete metric over $\mathbb{L}^t$. □

To define a distance over trace distribution formulae we see them as probability distribution over trace formulae and we define the distance over $\mathbb{L}^d$ as the Kantorovich lifting of the metric $\mathscr{D}_{\mathbb{L}}^{t}$.

Definition 27 (Distance on $\mathbb{L}^d$)**.** The function $\mathscr{D}_{\mathbb{L}}^{d}\colon \mathbb{L}^d \times \mathbb{L}^d \to [0,1]$ is defined over $\mathbb{L}^d$ as follows:

$$\mathscr{D}_{\mathbb{L}}^{d}(\Psi_1,\Psi_2) = \mathbf{K}(\mathscr{D}_{\mathbb{L}}^{t})(\Psi_1,\Psi_2).$$

Proposition 9. *The function $\mathscr{D}_{\mathbb{L}}^{d}$ is a 1-bounded metric over $\mathbb{L}^d$.*

Proof. First we prove that $\mathscr{D}_{\mathbb{L}}^{d}$ is a metric over $\mathbb{L}^d$, namely that

1. $\mathscr{D}_{\mathbb{L}}^{d}(\Psi_1,\Psi_2) = 0$ iff $\Psi_1 = \Psi_2$;
2. $\mathscr{D}_{\mathbb{L}}^{d}(\Psi_1,\Psi_2) = \mathscr{D}_{\mathbb{L}}^{d}(\Psi_2,\Psi_1)$;
3. $\mathscr{D}_{\mathbb{L}}^{d}(\Psi_1,\Psi_2) \le \mathscr{D}_{\mathbb{L}}^{d}(\Psi_1,\Psi_3) + \mathscr{D}_{\mathbb{L}}^{d}(\Psi_3,\Psi_2)$.

Proof of item 1

($\Leftarrow$) Assume first that $\Psi_1 = \Psi_2$. Then $\mathscr{D}_{\mathbb{L}}^{d}(\Psi_1,\Psi_2) = 0$ immediately follows from Definition 27, since the Kantorovich metric is a pseudometric.

($\Rightarrow$) Assume now that $\mathscr{D}_{\mathbb{L}}^{d}(\Psi_1,\Psi_2) = 0$. We aim to show that this implies that $\Psi_1 = \Psi_2$. Assume wlog. that $\Psi_1 = \bigoplus_{i\in I} r_i\Phi_i$ and that $\Psi_2 = \bigoplus_{j\in J} r_j\Phi_j$. Then we have

$$\mathscr{D}_{\mathbb{L}}^{d}(\Psi_1,\Psi_2) = \min_{\mathfrak{w}\in\mathfrak{W}(\Psi_1,\Psi_2)} \sum_{i\in I, j\in J} \mathfrak{w}(\Phi_i,\Phi_j)\mathscr{D}_{\mathbb{L}}^{t}(\Phi_i,\Phi_j) \tag{20}$$

and the distance in Equation (20) is 0 if, given the optimal matching $\tilde{\mathfrak{w}}$

$$\tilde{\mathfrak{w}}(\Phi_i,\Phi_j) > 0 \text{ iff } \mathscr{D}_{\mathbb{L}}^{t}(\Phi_i,\Phi_j) = 0.$$

By Proposition 8 we have that $\mathscr{D}_{\mathbb{L}}^{t}(\Phi_i,\Phi_j) = 0$ iff $\Phi_i = \Phi_j$. In particular, let Φ_{j_i} be any formula in $\{\Phi_j \mid j \in J\}$ s.t. $\Phi_i = \Phi_{j_i}$. Since by Definition 18 the trace formulae Φ_i occurring in Ψ_1 are pairwise distinct and, analogously, the trace formulae Φ_j occurring in Ψ_2 are pairwise distinct, we gather that

$$r_i = \textstyle\sum_{j\in J} \tilde{\mathfrak{w}}(\Phi_i,\Phi_j) = \sum_{j_i\in J} \tilde{\mathfrak{w}}(\Phi_i,\Phi_{j_i}) = \tilde{\mathfrak{w}}(\Phi_i,\Phi_{j_i})$$
$$r_j = \textstyle\sum_{i\in I} \tilde{\mathfrak{w}}(\Phi_i,\Phi_j) = \sum_{i_j\in I} \tilde{\mathfrak{w}}(\Phi_{i_j},\Phi_j) = \tilde{\mathfrak{w}}(\Phi_{i_j},\Phi_j).$$

Therefore we can infer that $\Psi_1 = \Psi_2$ as probability distributions over $\mathbb{L}^{\mathrm{t}}$.

Proof of item 2 Immediate from the discrete metric and the matching being both symmetric.

Proof of item 3 Assume wlog. that $\Psi_1 = \bigoplus_{i\in I} r_i\Phi_i$, $\Psi_2 = \bigoplus_{j\in J} r_j\Phi_j$ and $\Psi_3 = \bigoplus_{h\in H} r_h\Phi_h$.
Let $\mathfrak{w}_{1,3} \in \mathfrak{W}(\Psi_1,\Psi_3)$ be an optimal matching for Ψ_1,Ψ_3, namely

$$\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_1,\Psi_3) = \min_{\mathfrak{w}\in\mathfrak{W}(\Psi_1,\Psi_3)} \sum_{\substack{i\in I\\ h\in H}} \mathfrak{w}(\Phi_i,\Phi_h)\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_i,\Phi_h) = \sum_{\substack{i\in I\\ h\in H}} \mathfrak{w}_{1,3}(\Phi_i,\Phi_h)\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_i,\Phi_h)$$

and let $\mathfrak{w}_{2,3} \in \mathfrak{W}(\Psi_2,\Psi_3)$ be an optimal matching for Ψ_2,Ψ_3, that is

$$\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_2,\Psi_3) = \min_{\mathfrak{w}\in\mathfrak{W}(\Psi_2,\Psi_3)} \sum_{\substack{j\in J\\ h\in H}} \mathfrak{w}(\Phi_j,\Phi_h)\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_j,\Phi_h) = \sum_{\substack{j\in J\\ h\in H}} \mathfrak{w}_{2,3}(\Phi_j,\Phi_h)\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_j,\Phi_h).$$

Consider now the function $f\colon I\times J\times H \to [0,1]$ defined by

$$f(i,j,h) = \mathfrak{w}_{1,3}(\Phi_i,\Phi_h)\cdot\mathfrak{w}_{2,3}(\Phi_j,\Phi_h)\cdot\frac{1}{r_h}.$$

Then, we have $\sum_{j\in J} f(i,j,h) = \mathfrak{w}_{1,3}(\Phi_i,\Phi_h)$ namely the projection of f over the first and third components coincides with the optimal matching for Ψ_1,Ψ_3. Similarly, $\sum_{i\in I} f(i,j,h) = \mathfrak{w}_{2,3}(\Phi_j,\Phi_h)$ namely the projection of f over the second and third components coincides with the optimal matching for Ψ_2,Ψ_3. Moreover, it holds that $\sum_{j\in J,h\in H} f(i,j,h) = r_i$ and $\sum_{i\in I,h\in H} f(i,j,h) = r_j$, that is $f(i,j,h)$ is a matching in $\mathfrak{W}(\Psi_1,\Psi_2)$. Therefore,

$$\begin{aligned}
\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_1,\Psi_2) = {} & \min_{\mathfrak{w}\in\mathfrak{W}(\Psi_1,\Psi_2)} \textstyle\sum_{i\in I,\, j\in J} \mathfrak{w}(\Phi_i,\Phi_j)\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_i,\Phi_j) && \text{(by definition)}\\
\leq {} & \textstyle\sum_{i\in I,\, j\in J, h\in H} f(i,j,h)\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_i,\Phi_j) && \text{(by construction of } f\text{)}\\
\leq {} & \textstyle\sum_{i\in I,\, j\in J, h\in H} f(i,j,h)\big(\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_i,\Phi_h) + \mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_j,\Phi_h)\big) && \text{(since } \mathscr{D}^{\mathrm{t}}_{\mathbb{L}} \text{ is a metric)}\\
= {} & \textstyle\sum_{i\in I,\, j\in J, h\in H} f(i,j,h)\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_i,\Phi_h) + \\
& \textstyle\sum_{i\in I,\, j\in J, h\in H} f(i,j,h)\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_j,\Phi_h)\\
= {} & \textstyle\sum_{i\in I, h\in H} \Big(\sum_{j\in J} f(i,j,h)\Big)\cdot\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_i,\Phi_h) + \\
& \textstyle\sum_{j\in J, h\in H} \Big(\sum_{i\in I} f(i,j,h)\Big)\cdot\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_j,\Phi_h)\\
= {} & \textstyle\sum_{i\in I, h\in H} \mathfrak{w}_{1,3}(\Phi_i,\Phi_h)\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_i,\Phi_h) + \\
& \textstyle\sum_{j\in J, h\in H} \mathfrak{w}_{2,3}(\Phi_j,\Phi_h)\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_j,\Phi_h) && \text{(by construction of } f\text{)}\\
= {} & \mathbf{K}(\mathscr{D}^{\mathrm{t}}_{\mathbb{L}})(\Psi_1,\Psi_3) + \mathbf{K}(\mathscr{D}^{\mathrm{t}}_{\mathbb{L}})(\Psi_3,\Psi_2) && \text{(by definition of } \mathfrak{w}_{1,3},\mathfrak{w}_{2,3}\text{)}\\
= {} & \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_1,\Psi_3) + \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_3,\Psi_2) && \text{(by definition).}
\end{aligned}$$

To conclude, we need to show that $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}$ is 1-bounded, namely that for each $\Psi_1,\Psi_2 \in \mathbb{L}^{\mathrm{d}}$ we have $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_1,\Psi_2) \leq 1$. Assume wlog that $\Psi_1 = \bigoplus_{i\in I} r_i\Phi_i$ and $\Psi_2 = \bigoplus_{j\in J} r_j \in \Phi_j$. We have

$$\begin{aligned}
\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_1,\Psi_2) = {} & \min_{\mathfrak{w}\in\mathfrak{W}(\Psi_1,\Psi_2)} \textstyle\sum_{i\in I,j\in J} \mathfrak{w}(\Phi_i,\Phi_j)\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_i,\Phi_j)\\
\leq {} & \textstyle\sum_{i\in I,j\in J} \mathfrak{w}(\Phi_i,\Phi_j)\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_i,\Phi_j) && \text{(for an arbitrary } \mathfrak{w}\text{)}\\
\leq {} & \textstyle\sum_{i\in I,j\in J} \mathfrak{w}(\Phi_i,\Phi_j) && (\mathscr{D}^{\mathrm{t}}_{\mathbb{L}} \text{ is either 1 or 0})\\
= {} & 1 && (\mathfrak{w} \text{ is probability distribution}).
\end{aligned}$$

□

***Example* 6.** Consider the trace distribution formulae $\Psi_1 = 0.6\langle a\rangle\langle b\rangle\top \oplus 0.4\langle a\rangle\langle c\rangle\top$ and $\Psi_2 = 0.7\langle a\rangle\langle c\rangle\top \oplus 0.3\langle a\rangle\langle b\rangle\top$. We have that

$$\begin{aligned}
\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_1,\Psi_2) = \quad & \min_{\mathfrak{w}\in\mathfrak{W}(\Psi_1,\Psi_2)} \textstyle\sum_{\substack{\Phi\in\mathsf{supp}(\Psi_1)\\ \Phi'\in\mathsf{supp}(\Psi_2)}} \mathfrak{w}(\Phi,\Phi')\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi,\Phi') \\
\leq \quad & 0.3\cdot\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\langle a\rangle\langle b\rangle\top,\langle a\rangle\langle b\rangle\top) + 0.4\cdot\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\langle a\rangle\langle c\rangle\top,\langle a\rangle\langle c\rangle\top) + 0.3\cdot\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\langle a\rangle\langle b\rangle\top,\langle a\rangle\langle c\rangle\top) \\
= \quad & 0.3\cdot 0 + 0.4\cdot 0 + 0.3\cdot 1 \\
= \quad & 0.3
\end{aligned}$$

Next result derives from our characterization of trace distribution equivalence of resolutions (Theorem 2).

Theorem 6. *The kernel of $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}$ is trace distribution equivalence of resolutions.*

Proof. Let $s,t\in\mathbf{S}$ and consider $\mathscr{Z}_s\in\mathrm{Res}(s)$, with $z_s = \mathrm{corr}^{-1}_{\mathscr{Z}_s}(s)$, and $\mathscr{Z}_t\in\mathrm{Res}(t)$, with $z_t = \mathrm{corr}^{-1}_{\mathscr{Z}_s}(t)$. By Theorem 2 we have that $z_s \approx_{\mathrm{st}} z_t$ iff $\Psi_{\mathscr{Z}_s} = \Psi_{\mathscr{Z}_t}$. Since by Proposition 9 $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}$ is a metric on $\mathbb{L}^{\mathrm{d}}$, we have that $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_{\mathscr{Z}_s},\Psi_{\mathscr{Z}_t}) = 0$ iff $\Psi_{\mathscr{Z}_s} = \Psi_{\mathscr{Z}_t}$. Thus we can conclude that

$$z_s \approx_{\mathrm{st}} z_t \quad \text{iff} \quad \Psi_{\mathscr{Z}_s} = \Psi_{\mathscr{Z}_t} \quad \text{iff} \quad \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_{\mathscr{Z}_s},\Psi_{\mathscr{Z}_t}) = 0.$$

□

We lift the distance over formulae to a distance over processes as the Hausdorff distance between the sets of formulae satisfied by them.

Definition 28. The $\mathbb{L}$-*distance* over processes $\mathscr{D}_{\mathbb{L}}\colon \mathbf{S}\times\mathbf{S}\to[0,1]$ is defined, for all $s,t\in\mathbf{S}$, by

$$\mathscr{D}_{\mathbb{L}}(s,t) = \mathbf{H}(\mathscr{D}^{\mathrm{d}}_{\mathbb{L}})(\mathbb{L}(s),\mathbb{L}(t)).$$

Proposition 10. *The mapping $\mathscr{D}_{\mathbb{L}}$ is a* 1*-bounded pseudometric over* $\mathbf{S}$.

Proof. First we show that $\mathscr{D}_{\mathbb{L}}$ is a pseudometric over $\mathbf{S}$, namely that for each $s,t,u\in\mathbf{S}$

$$\mathscr{D}_{\mathbb{L}}(s,s) = 0 \tag{21}$$

$$\mathscr{D}_{\mathbb{L}}(s,t) = \mathscr{D}_{\mathbb{L}}(t,s) \tag{22}$$

$$\mathscr{D}_{\mathbb{L}}(s,t) \leq \mathscr{D}_{\mathbb{L}}(s,u) + \mathscr{D}_{\mathbb{L}}(u,t) \tag{23}$$

Equation (21) and Equation (22) are immediate from the definition of $\mathscr{D}_{\mathbb{L}}$ (Definition 28).

Let us prove Equation (23). Firstly, we notice that from the definition of Hausdorff distance we have

$$\mathscr{D}_{\mathbb{L}}(s,t) = \max\{\sup_{\Psi\in\mathbb{L}(s)}\inf_{\Psi'\in\mathbb{L}(t)} \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\Psi'),\ \sup_{\Psi'\in\mathbb{L}(t)}\inf_{\Psi\in\mathbb{L}(s)} \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\Psi')\}.$$

Thus, for all $s,t,u\in\mathbf{S}$ we can infer that

$$\sup_{\Psi\in\mathbb{L}(s)}\inf_{\Psi''\in\mathbb{L}(u)} \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\Psi'') \leq \mathscr{D}_{\mathbb{L}}(s,u) \tag{24}$$

$$\sup_{\Psi''\in\mathbb{L}(u)}\inf_{\Psi'\in\mathbb{L}(t)} \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi'',\Psi') \leq \mathscr{D}_{\mathbb{L}}(u,t). \tag{25}$$

As a first step, we aim to show that

$$\sup_{\Psi\in\mathbb{L}(s)}\inf_{\Psi'\in\mathbb{L}(t)} \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\Psi') \leq \mathscr{D}_{\mathbb{L}}(s,u) + \mathscr{D}_{\mathbb{L}}(u,t). \tag{26}$$

For sake of simplicity, we index formulae in $\mathbb{L}(s)$ by indexes in the set J, formulae in $\mathbb{L}(t)$ by indexes in set I and formulae in $\mathbb{L}(u)$ by indexes in H. By definition of infimum we have that for each $\varepsilon_1 > 0$

$$\text{for each } \Psi_j \in \mathbb{L}(s) \text{ there is a } \Psi_{h_j} \in \mathbb{L}(u) \text{ s.t. } \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_j, \Psi_{h_j}) < \inf_{\Psi_h \in \mathbb{L}(u)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_j, \Psi_h) + \varepsilon_1 \tag{27}$$

and analogously for each $\varepsilon_2 > 0$

$$\text{for each } \Psi_h \in \mathbb{L}(u) \text{ there is a } \Psi_{i_h} \in \mathbb{L}(t) \text{ s.t. } \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_h, \Psi_{i_h}) < \inf_{\Psi_i \in \mathbb{L}(t)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_h, \Psi_i) + \varepsilon_2. \tag{28}$$

In particular given $\Psi_j \in \mathbb{L}(s)$ let $\Psi_{h_j} \in \mathbb{L}(u)$ be the index realizing Equation (27), with respect to ε_1, and let $\Psi_{i_{h_j}} \in \mathbb{L}(t)$ be the index realizing Equation (28) with respect to Ψ_{h_j} and ε_2. Then we have

$$\begin{array}{lll} & \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_j, \Psi_{i_{h_j}}) & \\ \leq & \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_j, \Psi_{h_j}) + \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_{h_j}, \Psi_{i_{h_j}}) & (\mathscr{D}_{\mathbb{L}}^{\mathrm{d}} \text{ is a metric}) \\ < & \big(\inf_{\Psi_h \in \mathbb{L}(u)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_j, \Psi_h) + \varepsilon_1\big) + \big(\inf_{\Psi_i \in \mathbb{L}(t)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_{h_j}, \Psi_i) + \varepsilon_2\big) & \text{(by Eqs. 27 and 28)} \\ \leq & \big(\sup_{\Psi_j \in \mathbb{L}(s)} \inf_{\Psi_h \in \mathbb{L}(u)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_j, \Psi_h) + \varepsilon_1\big) + \big(\sup_{\Psi_h \in \mathbb{L}(u)} \inf_{\Psi_i \in \mathbb{L}(t)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_h, \Psi_i) + \varepsilon_2\big) & \end{array}$$

from which we gather

$$\inf_{\Psi_i \in \mathbb{L}(t)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_j, \Psi_i) \leq \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_j, \Psi_{i_{h_j}}) < \sup_{\Psi_j \in \mathbb{L}(s)} \inf_{\Psi_h \in \mathbb{L}(u)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_j, \Psi_h) + \sup_{\Psi_h \in \mathbb{L}(u)} \inf_{\Psi_i \in \mathbb{L}(t)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_h, \Psi_i) + \varepsilon_1 + \varepsilon_2.$$

Thus, since j was arbitrary, we obtain

$$\sup_{\Psi_j \in \mathbb{L}(s)} \inf_{\Psi_i \in \mathbb{L}(t)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_j, \Psi_i) \leq \sup_{\Psi_j \in \mathbb{L}(s)} \inf_{\Psi_h \in \mathbb{L}(u)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_j, \Psi_h) + \sup_{\Psi_h \in \mathbb{L}(u)} \inf_{\Psi_i \in \mathbb{L}(t)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_h, \Psi_i) + \varepsilon_1 + \varepsilon_2$$

and since this relation holds for any ε_1 and ε_2 we can conclude that

$$\sup_{\Psi_j \in \mathbb{L}(s)} \inf_{\Psi_i \in \mathbb{L}(t)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_j, \Psi_i) \leq \sup_{\Psi_j \in \mathbb{L}(s)} \inf_{\Psi_h \in \mathbb{L}(u)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_j, \Psi_h) + \sup_{\Psi_h \in \mathbb{L}(u)} \inf_{\Psi_i \in \mathbb{L}(t)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_h, \Psi_i).$$

Then, by the inequalities in Equation (24) and Equation (25) we can conclude that

$$\sup_{\Psi_j \in \mathbb{L}(s)} \inf_{\Psi_i \in \mathbb{L}(t)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_j, \Psi_i) \leq \mathscr{D}_{\mathbb{L}}(s,u) + \mathscr{D}_{\mathbb{L}}(u,t)$$

and thus Equation (26) holds. Switching the roles of s and t in the steps above allows us to infer

$$\sup_{\Psi_i \in \mathbb{L}(t)} \inf_{\Psi_j \in \mathbb{L}(s)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_j, \Psi_i) \leq \mathscr{D}_{\mathbb{L}}(s,u) + \mathscr{D}_{\mathbb{L}}(u,t). \tag{29}$$

Finally, we have

$$\begin{array}{rll} \mathscr{D}_{\mathbb{L}}(s,t) = & \max\{\sup_{\Psi_j \in \mathbb{L}(s)} \inf_{\Psi_i \in \mathbb{L}(t)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_j, \Psi_i), \sup_{\Psi_i \in \mathbb{L}(t)} \inf_{\Psi_j \in \mathbb{L}(s)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_j, \Psi_i)\} & \text{(by definition)} \\ \leq & \mathscr{D}_{\mathbb{L}}(s,u) + \mathscr{D}_{\mathbb{L}}(u,t) & \end{array}$$

where the last relation follows by Equations (26) and (29).

To conclude, we need to show that $\mathscr{D}_{\mathbb{L}}$ is 1-bounded. We have

$$\begin{array}{rll} \mathscr{D}_{\mathbb{L}}(s,t) = & \mathbf{H}(\mathscr{D}_{\mathbb{L}}^{\mathrm{d}})(\mathbb{L}(s), \mathbb{L}(t)) & \\ = & \max\Big\{\sup_{\Psi_i \in \mathbb{L}(s)} \inf_{\Psi_j \in \mathbb{L}(t)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_i, \Psi_j),\ \sup_{\Psi_j \in \mathbb{L}(t)} \inf_{\Psi_i \in \mathbb{L}(s)} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_i, \Psi_j)\Big\} & \\ \leq & \max\{1,\ 1\} & (\mathscr{D}_{\mathbb{L}}^{\mathrm{d}} \text{ is 1-bounded}) \\ = & 1. & \end{array}$$

□

Proposition 11. *Let $s \in \mathbf{S}$. The set $\mathbb{L}(s)$ is a closed subset of $\mathbb{L}$ wrt. the topology induced by $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}$.*

Proof. As we are working on a metric space, the proof obligation is equivalent to prove that each sequence in $\mathbb{L}(s)$ that admits a limit converges in $\mathbb{L}(s)$, namely

$$\text{for each } \{\Psi_n\}_{n\in\mathbb{N}} \subseteq \mathbb{L}(s) \text{ s.t. there is } \Psi \in \mathbb{L} \text{ with } \lim_{n\to\infty} \Psi_n = \Psi \text{ then } \Psi \in \mathbb{L}(s). \tag{30}$$

From Theorem 1 we have that $\mathbb{L}(s) = \{\top\} \cup \{\Psi_{\mathscr{Z}} \mid \mathscr{Z} \in \mathrm{Res}(s)\}$. Since a finite union of closed sets is closed, the proof obligation Equation (30) is equivalent to prove that

$$\{\top\} \text{ is closed} \tag{31}$$

$$\{\Psi_{\mathscr{Z}} \mid \mathscr{Z} \in \mathrm{Res}(s)\} \text{ is closed} \tag{32}$$

Equation (31) is immediate since the only sequence in $\{\top\}$ admitting a limit is the constant sequence $\Psi_n = \top$ for all $n \in \mathbb{N}$.

Let us deal now with Equation (32). First of all, we notice that sequences in $\{\Psi_{\mathscr{Z}} \mid \mathscr{Z} \in \mathrm{Res}(s)\}$ can be written in the general form

$$\Psi_n = \bigoplus_{i\in I_n} r_i^{(n)} \Phi_i^{(n)}$$

with $\{\bigoplus_{i\in I_n} r_i^{(n)} \Phi_i^{(n)}\}_{n\in\mathbb{N}} \subseteq \mathbb{L}(s) \setminus \{\top\}$.

Assume that there is a trace distribution formula $\Psi \in \mathbb{L}^{\mathrm{d}}$ s.t. $\lim_{n\to\infty} \Psi_n = \Psi$. We aim to show that $\Psi \in \mathbb{L}(s)$, namely that

$$\Psi = \Psi_{\mathscr{Z}} \text{ for some } \mathscr{Z} \in \mathrm{Res}(s). \tag{33}$$

In what follows, we assume wlog that limit trace distribution formula Ψ has the form $\Psi = \bigoplus_{j\in J} r_j \Phi_j$.

From $\{\bigoplus_{i\in I_n} r_i^{(n)} \Phi_i^{(n)}\}_{n\in\mathbb{N}} \subseteq \mathbb{L}(s) \setminus \{\top\}$ we gather that for each $n \in \mathbb{N}$ there is a resolution $\mathscr{Z}_n \in \mathrm{Res}(s)$ s.t. $\Psi_{\mathscr{Z}_n} = \bigoplus_{i\in I_n} r_i^{(n)} \Phi_i^{(n)}$. For each $n \in \mathbb{N}$, let $z_n = \mathrm{corr}^{-1}_{\mathscr{Z}_n}(s)$. Then $\Psi_{\mathscr{Z}_n} = \bigoplus_{i\in I_n} r_i^{(n)} \Phi_i^{(n)}$ implies that $I_n = \mathrm{Tr}(\mathscr{C}_{\max}(z_n))$, namely I_n is the set of traces to which the maximal computations of the process z_n are compatible. Hence, for each $i \in I_n$ we have that $\Phi_i^{(n)}$ is the tracing formula of trace i (or to be more formal of the trace indexed by i) and $r_i^{(n)} = \Pr(\mathscr{C}_{\max}(z_n, i))$.

We notice that

$$\begin{array}{ll} & \lim_{n\to\infty} \Psi_n = \Psi \\ \text{iff} & \lim_{n\to\infty} \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_n, \Psi) = 0 \\ \text{iff} & \lim_{n\to\infty} \mathbf{K}(\mathscr{D}^{\mathrm{t}}_{\mathbb{L}})(\Psi_n, \Psi) = 0 \end{array}$$

that is iff the sequence $\{\Psi_n\}_{n\in\mathbb{N}}$ converges to Ψ with respect to the Kantorovich metric. Since we are considering distributions with finite support, the convergence with respect to the Kantorovich metric is equivalent to the weak convergence of probability distributions (also called convergence in distribution) which states that $\lim_{n\to\infty} \Psi_n(\Phi) = \Psi(\Phi)$ for each continuity point $\Phi \in \mathbb{L}^{\mathrm{t}}$ of Ψ. Since the probability distribution over trace formulae Ψ is discrete and with finite support, its continuity points are the trace formulae which are not in its support. Hence, we have that $\lim_{n\to\infty} \Psi_n(\Phi) = 0$ for each $\Phi \notin \{\Phi_j \mid j \in J\}$. More specifically, we obtain that $\lim_{n\to\infty} I_n = J$ which gives that if there is an index $\tilde{\imath}$ s.t. $\lim_{n\to\infty} \Phi^{(n)}_{\tilde{\imath}} \notin \{\Phi_j \mid j \in J\}$, or if $\{\Phi^{(n)}_{\tilde{\imath}}\}_{n\in N}$ has no limit, then $\lim_{n\to\infty} r^{(n)}_{\tilde{\imath}} = 0$. Furthermore, since $\mathscr{D}^{\mathrm{t}}_{\mathbb{L}}$ is the discrete metric over $\mathbb{L}^{\mathrm{t}}$, we have that a sequence of trace formulae $\{\Phi^{(n)}\}_{n\in\mathbb{N}}$ converges to Φ iff the sequence is definitively constant, namely iff there is an $N \in \mathbb{N}$ s.t. $\Phi^{(n)} = \Phi$ for all $n \geq N$. Therefore, from $\lim_{n\to\infty} I_n = J$ we can infer that there is an $N \in \mathbb{N}$ s.t. $I_n = J$ for all $n \geq N$. Consequently, by construction of the sets

I_n, we obtain that $J = \mathrm{Tr}(\mathscr{C}_{\max}(z_N))$ thus giving that, for each $j \in J$, Φ_j is the tracing formula of the trace j (or more formally of the trace indexed by j) and $r_j = \Pr(\mathscr{C}_{\max}(z_N, j))$. Thus, from Definition 24, we infer that the resolution $\mathscr{Z}_N \in \mathrm{Res}(s)$, namely the resolution whose mimicking formula corresponds to the N-th trace distribution formula in the sequence $\{\Psi_n\}_{n\in\mathbb{N}}$, is s.t. $\Psi = \Psi_{\mathscr{Z}_N}$, thus proving Equation (33) and concluding the proof. □

From our $\mathbb{L}$-characterization of strong trace equivalence (Theorem 3) we obtain the following result.

Theorem 7. *The kernel of $\mathscr{D}_{\mathbb{L}}$ is trace equivalence.*

Proof. ($\Rightarrow$) Assume first that $s \approx_{\mathrm{st}} t$. We aim to show that this implies that $\mathscr{D}_{\mathbb{L}}(s,t) = 0$. By Theorem 3 we have that $s \approx_{\mathrm{st}} t$ implies that $\mathbb{L}(s) = \mathbb{L}(t)$ from which we gather

$$\mathscr{D}_{\mathbb{L}}(s,t) = \mathbf{H}(\mathscr{D}^{\mathrm{d}}_{\mathbb{L}})(\mathbb{L}(s), \mathbb{L}(t)) = 0.$$

($\Leftarrow$) Assume now that $\mathscr{D}_{\mathbb{L}}(s,t) = 0$. We aim to show that this implies that $s \approx_{\mathrm{st}} t$. Since $\mathbb{L}(s)$ and $\mathbb{L}(t)$ are closed by Proposition 11 and since $\mathscr{D}_{\mathbb{L}}$ is a pseudometric by Proposition 10, from $\mathscr{D}_{\mathbb{L}}(s,t) = 0$ we can infer that $\mathbb{L}(s) = \mathbb{L}(t)$. By Theorem 3 we can conclude that $s \approx_{\mathrm{st}} t$. □

Finally, we obtain the characterization of the strong trace metric.

Theorem 8 (Characterization of strong trace metric)**.** *For all $s,t \in \mathbf{S}$ we have $\mathbf{d}_T(s,t) = \mathscr{D}_{\mathbb{L}}(s,t)$.*

Proof. By definition of trace metric (Definition 14) we have that

$$\mathbf{d}_T(s,t) = \max\left\{ \sup_{\mathscr{Z}_s \in \mathrm{Res}(s)} \inf_{\mathscr{Z}_t \in \mathrm{Res}(t)} D_T(\mathscr{Z}_s, \mathscr{Z}_t),\ \sup_{\mathscr{Z}_t \in \mathrm{Res}(t)} \inf_{\mathscr{Z}_s \in \mathrm{Res}(s)} D_T(\mathscr{Z}_s, \mathscr{Z}_t) \right\}.$$

By definition of $\mathbb{L}$-distance over processes (Definition 28) we have that

$$\begin{aligned}
\mathscr{D}_{\mathbb{L}}(s,t) = {} & \mathbf{H}(\mathscr{D}^{\mathrm{d}}_{\mathbb{L}})(\mathbb{L}(s), \mathbb{L}(t)) \\
= {} & \mathbf{H}(\mathscr{D}^{\mathrm{d}}_{\mathbb{L}})(\{\top\} \cup \{\Psi_{\mathscr{Z}_s} \mid \mathscr{Z}_s \in \mathrm{Res}(s)\}, \{\top\} \cup \{\Psi_{\mathscr{Z}_t} \mid \mathscr{Z}_t \in \mathrm{Res}(t)\}) \\
= {} & \mathbf{H}(\mathscr{D}^{\mathrm{d}}_{\mathbb{L}})(\{\Psi_{\mathscr{Z}_s} \mid \mathscr{Z}_s \in \mathrm{Res}(s)\}, \{\Psi_{\mathscr{Z}_t} \mid \mathscr{Z}_t \in \mathrm{Res}(t)\}) \\
= {} & \max\left\{ \sup_{\mathscr{Z}_s \in \mathrm{Res}(s)} \inf_{\mathscr{Z}_t \in \mathrm{Res}(t)} \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_{\mathscr{Z}_s}, \Psi_{\mathscr{Z}_t}),\ \sup_{\mathscr{Z}_t \in \mathrm{Res}(t)} \inf_{\mathscr{Z}_s \in \mathrm{Res}(s)} \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_{\mathscr{Z}_s}, \Psi_{\mathscr{Z}_t}) \right\}
\end{aligned}$$

where the third equality follows from the fact that by Definition 27 we have $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\top, \top) = 0$ and $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\top, \Psi) = 1$ for any $\Psi \neq \top$. Thus we have that $\top = \mathrm{argmin}_{\Psi \in \{\top\} \cup \{\Psi_{\mathscr{Z}_t} \mid \mathscr{Z}_t \in \mathrm{Res}(t)\}} \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\top, \Psi)$ and symmetrically $\top = \mathrm{argmin}_{\Psi \in \{\top\} \cup \{\Psi_{\mathscr{Z}_s} \mid \mathscr{Z}_s \in \mathrm{Res}(s)\}} \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi, \top)$. Moreover, for any $\Psi \neq \top$ we have that $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi, \Psi') \leq \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi, \top)$ for any $\Psi' \in \{\Psi_{\mathscr{Z}_t} \mid \mathscr{Z}_t \in \mathrm{Res}(t)\}$ and $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi'', \Psi) \leq \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\top, \Psi)$ for any $\Psi'' \in \{\Psi_{\mathscr{Z}_s} \mid \mathscr{Z}_s \in \mathrm{Res}(s)\}$.

Hence, to prove the thesis it is enough to show that

$$D_T(\mathscr{Z}_s, \mathscr{Z}_t) = \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_{\mathscr{Z}_s}, \Psi_{\mathscr{Z}_t}) \text{ for all } \mathscr{Z}_s \in \mathrm{Res}(s), \mathscr{Z}_t \in \mathrm{Res}(t). \tag{34}$$

Let $\mathscr{Z}_s \in \mathrm{Res}(s)$, with $z_s = \mathrm{corr}^{-1}_{\mathscr{Z}_s}(s)$, and $\mathscr{Z}_t \in \mathrm{Res}(t)$, with $z_t = \mathrm{corr}^{-1}_{\mathscr{Z}_t}(t)$. Then by definition of mimicking formula (Definition 24) we have

$$\Psi_{\mathscr{Z}_s} = \bigoplus_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z_s))} \Pr(\mathscr{C}_{\max}(z_s, \alpha)) \Phi_\alpha$$

where for each $\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z_s))$ we have that Φ_α is the tracing formula for the trace α. Similarly,

$$\Psi_{\mathscr{Z}_t} = \bigoplus_{\beta \in \mathrm{Tr}(\mathscr{C}_{\max}(z_t))} \mathrm{Pr}(\mathscr{C}_{\max}(z_t,\beta))\Phi_\beta$$

where for each $\beta \in \mathrm{Tr}(\mathscr{C}_{\max}(z_t))$ we have that Φ_β is the tracing formula for the trace β.

By definition of trace distance between resolutions (Definition 13) we have that

$$D_T(\mathscr{Z}_s,\mathscr{Z}_t) = \min_{\mathfrak{w} \in \mathfrak{W}(\mathscr{T}_{\mathscr{Z}_s},\mathscr{T}_{\mathscr{Z}_t})} \sum_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z_s)), \beta \in \mathrm{Tr}(\mathscr{C}_{\max}(z_t))} \mathfrak{w}(\alpha,\beta) d_T(\alpha,\beta) \tag{35}$$

where, by definition of trace distance between traces (Definition 11), we have that $d_t(\alpha,\beta) = 0$ if $\alpha = \beta$ and $d_t(\alpha,\beta) = 1$ otherwise.

Hence, by definition of tracing formula (Definition 23), it is immediate that for all $\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z_s)), \beta \in \mathrm{Tr}(\mathscr{C}_{\max}(z_t))$ we have $d_T(\alpha,\beta) = \mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_\alpha,\Phi_\beta)$, thus giving

$$(35) = \min_{\mathfrak{w} \in \mathfrak{W}(\mathscr{T}_{\mathscr{Z}_s},\mathscr{T}_{\mathscr{Z}_t})} \sum_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z_s)), \beta \in \mathrm{Tr}(\mathscr{C}_{\max}(z_t))} \mathfrak{w}(\alpha,\beta) \mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_\alpha,\Phi_\beta). \tag{36}$$

Let $\tilde{\mathfrak{w}}$ be an optimal matching for $D_T(\mathscr{Z}_s,\mathscr{Z}_t)$, namely

$$(36) = \sum_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z_s)), \beta \in \mathrm{Tr}(\mathscr{C}_{\max}(z_t))} \tilde{\mathfrak{w}}(\alpha,\beta) \mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_\alpha,\Phi_\beta). \tag{37}$$

Then, by definition of matching and of $\mathscr{T}_{\_}$ (Definition 12) we have that for any $\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z_s)), \beta \in \mathrm{Tr}(\mathscr{C}_{\max}(z_t))$

$$\mathrm{Pr}(\mathscr{C}_{\max}(z_s,\alpha)) = \mathscr{T}_{\mathscr{Z}_s}(\alpha) = \textstyle\sum_{\beta \in \mathrm{Tr}(\mathscr{C}_{\max}(z_t))} \tilde{\mathfrak{w}}(\alpha,\beta)$$
$$\mathrm{Pr}(\mathscr{C}_{\max}(z_t,\beta)) = \mathscr{T}_{\mathscr{Z}_t}(\beta) = \textstyle\sum_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z_s))} \tilde{\mathfrak{w}}(\alpha,\beta).$$

Therefore we have obtained that $\tilde{\mathfrak{w}}$ is a matching for $\Psi_{\mathscr{Z}_s}$ and $\Psi_{\mathscr{Z}_t}$. In particular we notice that $\tilde{\mathfrak{w}}$ is actually an optimal matching for $\Psi_{\mathscr{Z}_s},\Psi_{\mathscr{Z}_t}$. This follows from the optimality of $\tilde{\mathfrak{w}}$ for $\mathscr{T}_{\mathscr{Z}_s},\mathscr{T}_{\mathscr{Z}_t}$. In fact each matching for $\Psi_{\mathscr{Z}_s},\Psi_{\mathscr{Z}_t}$ can be constructed from a matching for $\mathscr{T}_{\mathscr{Z}_s},\mathscr{T}_{\mathscr{Z}_t}$ using the same technique proposed above. Moreover, given $\mathfrak{w}_1 \in \mathfrak{W}(\mathscr{T}_{\mathscr{Z}_s},\mathscr{T}_{\mathscr{Z}_t})$ and $\mathfrak{w}_2$ being the matching for Ψ_1,Ψ_2 built from it, the reasoning above guarantees that

$$\sum_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z_s)), \beta \in \mathrm{Tr}(\mathscr{C}_{\max}(z_t))} \mathfrak{w}_1(\alpha,\beta) d_T(\alpha,\beta) = \sum_{\substack{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z_s)) \\ \beta \in \mathrm{Tr}(\mathscr{C}_{\max}(z_t))}} \mathfrak{w}_2(\alpha,\beta) \mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_\alpha,\Phi_\beta).$$

$\tilde{\mathfrak{w}}$ being optimal for D_T implies $\tilde{\mathfrak{w}}$ being optimal for $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}$. Hence by Definition 27 we have

$$\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_{\mathscr{Z}_s},\Psi_{\mathscr{Z}_t}) = \sum_{\alpha \in \mathrm{Tr}(\mathscr{C}_{\max}(z_s)), \beta \in \mathrm{Tr}(\mathscr{C}_{\max}(z_t))} \tilde{\mathfrak{w}}(\alpha,\beta) \mathscr{D}^{\mathrm{t}}_{\mathbb{L}}(\Phi_\alpha,\Phi_\beta).$$

From Equation (37) we infer $D_T(\mathscr{Z}_s,\mathscr{Z}_t) = \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_{\mathscr{Z}_s},\Psi_{\mathscr{Z}_t})$ thus proving Equation (34) and concluding the proof. □

6.2 $\mathbb{L}_{\mathrm{w}}$-characterization of weak trace metric

The idea behind the definition of a metric on $\mathbb{L}_{\mathrm{w}}$ is pretty much the same to the strong case. The main difference is that the distance on $\mathbb{L}_{\mathrm{w}}$ is a pseudometric whose kernel is given by $\mathbb{L}_{\mathrm{w}}$-equivalence.

Definition 29 (Distance on $\mathbb{L}_{\mathrm{w}}^{\mathrm{t}}$)**.** The function $\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}^{\mathrm{t}} \colon \mathbb{L}_{\mathrm{w}}^{\mathrm{t}} \times \mathbb{L}_{\mathrm{w}}^{\mathrm{t}} \to [0,1]$ is defined over $\mathbb{L}_{\mathrm{w}}^{\mathrm{t}}$ as follows:

$$\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}^{\mathrm{t}}(\Phi_1,\Phi_2) = \begin{cases} 0 & \text{if } \Phi_1 \equiv_{\mathrm{w}} \Phi_2 \\ 1 & \text{otherwise.} \end{cases}$$

Clearly, $\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}^{\mathrm{t}}$ is a pseudometric on $\mathbb{L}_{\mathrm{w}}^{\mathrm{t}}$ whose kernel is given by equivalence of trace formulae and we can lift it to a pseudometric over $\mathbb{L}_{\mathrm{w}}^{\mathrm{d}}$ via the Kantorovich lifting functional.

Definition 30 (Distance on $\mathbb{L}_{\mathrm{w}}^{\mathrm{d}}$)**.** The function $\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}^{\mathrm{d}} \colon \mathbb{L}_{\mathrm{w}}^{\mathrm{d}} \times \mathbb{L}_{\mathrm{w}}^{\mathrm{d}} \to [0,1]$ is defined over $\mathbb{L}_{\mathrm{w}}^{\mathrm{d}}$ as follows:

$$\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}^{\mathrm{d}}(\Psi_1,\Psi_2) = \mathbf{K}(\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}^{\mathrm{t}})(\Psi_1,\Psi_2).$$

Proposition 12. *The function $\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}^{\mathrm{d}}$ is a 1-bounded pseudometric over $\mathbb{L}_{\mathrm{w}}^{\mathrm{d}}$.*

Proof. The same arguments used in the proof of Proposition 9 apply, where in place of item 1 we simply need to show that $\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}^{\mathrm{d}}(\Psi,\Psi) = 0$, which is immediate from the definition through the Kantorovich pseudometric. □

Theorem 9. *The kernel of $\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}^{\mathrm{d}}$ is $\mathbb{L}_{\mathrm{w}}$-equivalence of trace distribution formulae.*

Proof. ($\Rightarrow$) Assume first that $\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}^{\mathrm{d}}(\Psi_1,\Psi_2) = 0$ for $\Psi_1 = \bigoplus_{i\in I} r_i\Phi_i$ and $\Psi_2 = \bigoplus_{j\in J} r_j\Phi_j$. We aim to show that this implies $\Psi_1 \equiv_{\mathrm{w}}^{\dagger} \Psi_2$. From the assumption, we have

$$\begin{aligned} 0 = \;& \mathscr{D}_{\mathbb{L}_{\mathrm{w}}}^{\mathrm{d}}(\textstyle\bigoplus_{i\in I} r_i\Phi_i, \bigoplus_{j\in J} r_j\Phi_j) \\ = \;& \min_{\mathfrak{w}\in\mathfrak{W}(\Psi_1,\Psi_2)} \textstyle\sum_{i\in I,\, j\in J} \mathfrak{w}(\Phi_i,\Phi_j)\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}^{\mathrm{t}}(\Phi_i,\Phi_j) \\ = \;& \textstyle\sum_{i\in I,\, j\in J} \mathfrak{w}(\Phi_i,\Phi_j)\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}^{\mathrm{t}}(\Phi_i,\Phi_j) && \text{(for } \mathfrak{w} \text{ optimal matching).} \end{aligned}$$

Thus, for each $i \in I$ and $j \in J$ we can distinguish two cases:

- either $\mathfrak{w}(\Phi_i,\Phi_j) = 0$,
- or $\mathfrak{w}(\Phi_i,\Phi_j) > 0$, implying $\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}^{\mathrm{t}}(\Phi_i,\Phi_j) = 0$, which is equivalent to say that $\Phi_i \equiv_{\mathrm{w}} \Phi_j$ by Definition 29.

For each $i \in I$, let $J_i \subseteq J$ be the set of indexes j_i for which $\mathfrak{w}(\Phi_i,\Phi_{j_i}) > 0$ and, symmetrically, for each $j \in J$ let $I_j \subseteq I$ be the set of indexes i_j for which $\mathfrak{w}(\Phi_{i_j},\Phi_j) > 0$. So we have

$$\begin{aligned} \Psi_1 = \;& \textstyle\bigoplus_{i\in I} r_i\Phi_i \\ = \;& \textstyle\bigoplus_{i\in I} \big(\sum_{j\in J} \mathfrak{w}(\Phi_i,\Phi_j)\big)\Phi_i && (\mathfrak{w} \in \mathfrak{W}(\Psi_1,\Psi_2)) \\ \equiv_{\mathrm{w}}^{\dagger} \;& \textstyle\bigoplus_{i\in I} \big(\sum_{j_i\in J_i} \mathfrak{w}(\Phi_i,\Phi_{j_j})\big)\Phi_i && \text{(by construction of each } J_i) \\ \equiv_{\mathrm{w}}^{\dagger} \;& \textstyle\bigoplus_{i\in I,\, j_i\in J_i} \mathfrak{w}(\Phi_i,\Phi_{j_i})\Phi_{j_i} && (\Phi_i \equiv_{\mathrm{w}} \Phi_{j_i} \text{ for each } j_i \in J_i) \\ \equiv_{\mathrm{w}}^{\dagger} \;& \textstyle\bigoplus_{i\in I,\, j_i\in J_i,\, i'_{j_i}\in I_{j_i}} \mathfrak{w}(\Phi_{i'_{j_i}},\Phi_{j_i})\Phi_{i'_{j_i}} && (\Phi_{i'_{j_i}} \equiv_{\mathrm{w}} \Phi_{j_i} \text{ for each } i'_{j_i} \in I_{j_i}) \\ \equiv_{\mathrm{w}}^{\dagger} \;& \textstyle\bigoplus_{i_j\in I_j,\, j\in J} \mathfrak{w}(\Phi_{i_j},\Phi_j)\Phi_{i_j} && \text{(all indexes } j \in J \text{ are involved)} \\ \equiv_{\mathrm{w}}^{\dagger} \;& \textstyle\bigoplus_{j\in J} \big(\sum_{i_j\in I_j} \mathfrak{w}(\Phi_{i_j},\Phi_j)\big)\Phi_j && (\Phi_j \equiv_{\mathrm{w}} \Phi_{i_j} \text{ for each } i_j \in I_j) \\ \equiv_{\mathrm{w}}^{\dagger} \;& \textstyle\bigoplus_{j\in J} \big(\sum_{i\in I} \mathfrak{w}(\Phi_i,\Phi_j)\big)\Phi_j && \text{(by construction of each } I_j) \\ = \;& \textstyle\bigoplus_{j\in J} r_j\Phi_j && (\mathfrak{w} \in \mathfrak{W}(\Psi_1,\Psi_2)) \\ = \;& \Psi_2. \end{aligned}$$

($\Leftarrow$). Assume that $\Psi_1 \equiv_{\mathrm{w}}^{\dagger} \Psi_2$. We aim to show that $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}_{\mathrm{w}}}(\Psi_1,\Psi_2)=0$. Assume wlog. that $\Psi_1 = \bigoplus_{i\in I} r_i\Phi_i$. By definition of $\equiv_{\mathrm{w}}$ (Definition 7) and definition of lifting of a relation (Definition 2), from $\Psi_2 \equiv_{\mathrm{w}}^{\dagger} \bigoplus_{i\in I} r_i\Phi_i$ we gather $\Psi_2 = \bigoplus_{\substack{i\in I\\ j_i\in J_i}} r_{j_i}\Phi_{j_i}$ with $\sum_{j_i\in J_i} r_{j_i} = r_i$ and $\Phi_{j_i} \equiv_{\mathrm{w}} \Phi_i$ for all $j_i \in J_i, i\in I$. Then

$$\begin{array}{rll}
\mathscr{D}^{\mathrm{d}}_{\mathbb{L}_{\mathrm{w}}}(\Psi_1,\Psi_2) = & \mathscr{D}^{\mathrm{d}}_{\mathbb{L}_{\mathrm{w}}}(\bigoplus_{i\in I} r_i\Phi_i, \bigoplus_{i\in I,\, j_i\in J_i} r_{j_i}\Phi_{j_i}) & \\
= & \min_{\mathfrak{w}\in\mathfrak{W}(\Psi_1,\Psi_2)} \sum_{\substack{i\in I,\, j_h\in J_h\\ h\in I}} \mathfrak{w}(\Phi_i,\Phi_{j_h}) \mathscr{D}^{\mathrm{t}}_{\mathbb{L}_{\mathrm{w}}}(\Phi_i,\Phi_{j_h}) & \\
\leq & \sum_{\substack{i\in I,\, j_h\in J_h\\ h\in I}} \tilde{\mathfrak{w}}(\Phi_i,\Phi_{j_h}) \mathscr{D}^{\mathrm{t}}_{\mathbb{L}_{\mathrm{w}}}(\Phi_i,\Phi_{j_h}) & \\
= & \sum_{i\in I,\, j_i\in J_i} r_{j_i} \mathscr{D}^{\mathrm{t}}_{\mathbb{L}_{\mathrm{w}}}(\Phi_i,\Phi_{j_i}) & \\
= & 0 & (\Phi_i \equiv_{\mathrm{w}} \Phi_{j_i} \text{ for each } j_i\in J_i \text{ and Def. 29})
\end{array}$$

where the inequality follows by observing that function $\tilde{\mathfrak{w}}$ defined by $\tilde{\mathfrak{w}}(\Phi_i,\Phi_{j_h}) = r_{j_i}$ if $h=i$ and $\tilde{\mathfrak{w}}(\Phi_i,\Phi_{j_h}) = 0$ otherwise, is a matching in $\mathfrak{W}(\Psi_1,\Psi_2)$. ☐

Corollary 1. *$\mathscr{Z}_1,\mathscr{Z}_2 \in \mathrm{Res}(\mathbf{S})$ are weak trace distribution equivalent iff $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}_{\mathrm{w}}}(\Psi_{\mathscr{Z}_1},\Psi_{\mathscr{Z}_2})=0$.*

Proof. ($\Rightarrow$) Assume first that $\mathscr{Z}_1$ and $\mathscr{Z}_2$ are weak trace distribution equivalent. Then from Theorem 4 we infer that $\Psi_{\mathscr{Z}_1} \equiv_{\mathrm{w}} \Psi_{\mathscr{Z}_2}$. By Theorem 9 this implies $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}_{\mathrm{w}}}(\Psi_{\mathscr{Z}_1},\Psi_{\mathscr{Z}_2})=0$.

($\Leftarrow$) Assume now that $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}_{\mathrm{w}}}(\Psi_{\mathscr{Z}_1},\Psi_{\mathscr{Z}_2})=0$. Then from Theorem 9 we infer that $\Psi_{\mathscr{Z}_1} \equiv_{\mathrm{w}} \Psi_{\mathscr{Z}_2}$. By Theorem 4 this implies that $\mathscr{Z}_1$ and $\mathscr{Z}_2$ are weak trace distribution equivalent. ☐

By the Hausdorff functional we lift the pseudometric $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}_{\mathrm{w}}}$ to a pseudometric over processes.

Definition 31. The $\mathbb{L}_{\mathrm{w}}$-*distance* over processes $\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}\colon \mathbf{S}\times\mathbf{S}\to[0,1]$ is defined, for all $s,t\in\mathbf{S}$, by

$$\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}(s,t) = \mathbf{H}(\mathscr{D}^{\mathrm{d}}_{\mathbb{L}_{\mathrm{w}}})(\mathbb{L}_{\mathrm{w}}(s),\mathbb{L}_{\mathrm{w}}(t)).$$

Proposition 13. *The mapping $\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}$ is a 1-bounded pseudometric over* $\mathbf{S}$.

Proof. The same arguments used in the proof of Proposition 10 apply. ☐

Proposition 14. *Let $s\in\mathbf{S}$. The set $\mathbb{L}_{\mathrm{w}}(s)$ is a closed subset of $\mathbb{L}_{\mathrm{w}}$ wrt. the topology induced by $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}_{\mathrm{w}}}$.*

Proof. Since $(\mathbb{L}^{\mathrm{d}}_{\mathrm{w}},\mathscr{D}^{\mathrm{d}}_{\mathbb{L}_{\mathrm{w}}})$ is a pseudometric space (Proposition 12 and Theorem 9), to prove the thesis we need to show that the quotient space $\mathbb{L}_{\mathrm{w}}(s)_{/\equiv_{\mathrm{w}}}$ is a closed subset of $\mathbb{L}_{\mathrm{w}/\equiv_{\mathrm{w}}}$ with respect to the topology induced by $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}_{\mathrm{w}}}$ (in fact $(\mathbb{L}^{\mathrm{d}}_{\mathrm{w}/\equiv_{\mathrm{w}}},\mathscr{D}^{\mathrm{d}}_{\mathbb{L}_{\mathrm{w}}})$ is a metric space). From Remark 2 we have that $\mathbb{L}^{\mathrm{d}}_{\mathrm{w}/\equiv_{\mathrm{w}}} = \mathbb{L}^{\mathrm{d}}$ and $\mathbb{L}_{\mathrm{w}}(s)_{/\equiv_{\mathrm{w}}} = \mathbb{L}(s)$. Moreover, we have that $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}_{\mathrm{w}}}|_{\mathbb{L}^{\mathrm{d}}_{\mathrm{w}/\equiv_{\mathrm{w}}}} = \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}$. Hence, the same arguments used in the proof of Proposition 11 allow us to prove that $\mathbb{L}_{\mathrm{w}}(s)_{/\equiv_{\mathrm{w}}}$ is a closed subset of $\mathbb{L}_{\mathrm{w}/\equiv_{\mathrm{w}}}$ wrt. the topology induced by $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}_{\mathrm{w}}}$. This gives the result also for $\mathbb{L}_{\mathrm{w}}(s)$ wrt to $\mathbb{L}_{\mathrm{w}}$ and $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}_{\mathrm{w}}}$. ☐

Theorem 10. *The kernel of $\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}$ is weak trace equivalence.*

Proof. ($\Rightarrow$) Assume that $s\approx_{\mathrm{wt}} t$. We aim to show that $\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}(s,t)=0$. By Theorem 5 we have that $s\approx_{\mathrm{wt}} t$ implies that $\mathbb{L}_{\mathrm{w}}(s) \equiv_{\mathrm{w}}^{\dagger} \mathbb{L}_{\mathrm{w}}(t)$. Since the kernel of $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}_{\mathrm{w}}}$ is given by $\equiv_{\mathrm{w}}^{\dagger}$ (Theorem 9), we can infer

$$\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}(s,t) = \mathbf{H}(\mathscr{D}^{\mathrm{t}}_{\mathbb{L}_{\mathrm{w}}})(\mathbb{L}_{\mathrm{w}}(s),\mathbb{L}_{\mathrm{w}}(t)) = 0.$$

($\Leftarrow$) Assume now that $\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}(s,t)=0$. We aim to show that this implies that $s\approx_{\mathrm{wt}} t$. Since (i) $\mathbb{L}_{\mathrm{w}}(s)$ and $\mathbb{L}_{\mathrm{w}}(t)$ are closed by Proposition 14, (ii) $\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}$ is a pseudometric by Proposition 13 and (iii) the kernel of $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}_{\mathrm{w}}}$ is $\equiv_{\mathrm{w}}^{\dagger}$ by Theorem 9, from $\mathscr{D}_{\mathbb{L}_{\mathrm{w}}}(s,t)=0$ we can infer $\mathbb{L}_{\mathrm{w}}(s) \equiv_{\mathrm{w}}^{\dagger} \mathbb{L}_{\mathrm{w}}(t)$. Then, by Theorem 5 we can conclude $s\approx_{\mathrm{wt}} t$. ☐

Finally, we obtain the characterization of the weak trace metric.

Theorem 11 (Characterization of weak trace metric)**.** *For all* $s,t \in \mathbf{S}$ *we have* $\mathbf{d}_T^{\mathrm{w}}(s,t) = \mathscr{D}_{\mathbb{L}_{\mathrm{w}}}(s,t)$.

Proof. The same arguments used in the proof of Thm 8 apply. □

7 From boolean to real semantics

In this section we focus on $\mathbb{L}$ and we exploit the distance between formulae to define a real valued semantics for it, namely given a process s we assign to each formula a value in $[0,1]$ expressing the probability that s satisfies it. Then we show that our logical characterization of trace metric can be restated in terms of the general schema $\mathbf{d}_T(s,t) = \sup_{\Psi \in \mathbb{L}^{\mathrm{d}}} |\,[\Psi](s) - [\Psi](t)\,|$ where $[\Psi](s)$ denotes the value of the formula Ψ at process s, accordingly to the new real valued semantics. We remark that although, due to space restrictions, we present only the result for $\mathbb{L}$, the technique we propose would lead to the same results when applied to $\mathbb{L}_{\mathrm{w}}$.

First of all, we recall the notion of *distance function*, namely the distance between a point and a set.

Definition 32 (Distance function)**.** Let $\mathbb{L}' \subseteq \mathbb{L}^{\mathrm{d}}$. Given any $\Psi \in \mathbb{L}^{\mathrm{d}}$ we denote by $\mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi,\mathbb{L}')$ the *distance between* Ψ *and the set* $\mathbb{L}'$ defined by $\mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi,\mathbb{L}') = \inf_{\Psi' \in \mathbb{L}'} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi,\Psi')$.

Then we obtain the following characterization of the Hausdorff distance.

Proposition 15. *Let* $\mathbb{L}_1,\mathbb{L}_2 \subseteq \mathbb{L}^{\mathrm{d}}$. *Then it holds that* $\mathbf{H}(\mathscr{D}_{\mathbb{L}}^{\mathrm{d}})(\mathbb{L}_1,\mathbb{L}_2) = \sup_{\Psi \in \mathbb{L}^{\mathrm{d}}} |\mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi,\mathbb{L}_1) - \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi,\mathbb{L}_2)|$.

Proof. It is clear that

$$\mathbf{H}(\mathscr{D}_{\mathbb{L}}^{\mathrm{d}})(\mathbb{L}_1,\mathbb{L}_2) = \max\left\{ \sup_{\Psi_1 \in \mathbb{L}_1} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_1,\mathbb{L}_2),\ \sup_{\Psi_2 \in \mathbb{L}_2} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_2,\mathbb{L}_1) \right\}. \tag{38}$$

Firstly we show that

$$\mathbf{H}(\mathscr{D}_{\mathbb{L}}^{\mathrm{d}})(\mathbb{L}_1,\mathbb{L}_2) \le \sup_{\Psi \in \mathbb{L}^{\mathrm{d}}} |\mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi,\mathbb{L}_1) - \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi,\mathbb{L}_2)|. \tag{39}$$

Without loss of generality, we can assume that $\mathbf{H}(\mathscr{D}_{\mathbb{L}}^{\mathrm{d}})(\mathbb{L}_1,\mathbb{L}_2) = \sup_{\Psi_1 \in \mathbb{L}_1} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_1,\mathbb{L}_2)$. Then we have

$$\begin{aligned} \sup_{\Psi_1 \in \mathbb{L}_1} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_1,\mathbb{L}_2) = &\quad \sup_{\Psi_1 \in \mathbb{L}_1} |\mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_1,\mathbb{L}_2) - \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_1,\mathbb{L}_1)| \\ \le &\quad \sup_{\Psi \in \mathbb{L}^{\mathrm{d}}} |\mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi,\mathbb{L}_2) - \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi,\mathbb{L}_1)| \end{aligned}$$

from which Equation (39) holds.

Next, we aim to show the converse inequality, namely

$$\mathbf{H}(\mathscr{D}_{\mathbb{L}}^{\mathrm{d}})(\mathbb{L}_1,\mathbb{L}_2) \ge \sup_{\Psi \in \mathbb{L}^{\mathrm{d}}} |\mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi,\mathbb{L}_1) - \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi,\mathbb{L}_2)|. \tag{40}$$

To this aim, we show that

$$\text{for each } \Psi \in \mathbb{L}^{\mathrm{d}} \text{ it holds } |\mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi,\mathbb{L}_1) - \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi,\mathbb{L}_2)| \le \mathbf{H}(\mathscr{D}_{\mathbb{L}}^{\mathrm{d}})(\mathbb{L}_1,\mathbb{L}_2). \tag{41}$$

- Assume $\Psi \in \mathbb{L}_1$. Then $\mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi,\mathbb{L}_1) = 0$ so that $|\mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi,\mathbb{L}_1) - \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi,\mathbb{L}_2)| = \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi,\mathbb{L}_2)$. Moreover

  $$\mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi,\mathbb{L}_2) \le \sup_{\Psi_1 \in \mathbb{L}_1} \mathscr{D}_{\mathbb{L}}^{\mathrm{d}}(\Psi_1,\mathbb{L}_2) \le \mathbf{H}(\mathscr{D}_{\mathbb{L}}^{\mathrm{d}})(\mathbb{L}_1,\mathbb{L}_2)$$

  and Equation (41) follows in this case.

- The case of $\Psi \in \mathbb{L}_2$ is analogous and therefore Equation (41) follows also in this case.
- Finally, assume that $\Psi \notin \mathbb{L}_1 \cup \mathbb{L}_2$. Without loss of generality, we can assume that $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}_1) \geq \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}_2)$. By definition of *infimum* it holds that for each $\varepsilon > 0$ there is a formula $\Psi_\varepsilon \in \mathbb{L}_2$ s.t.

$$\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\Psi_\varepsilon) < \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}_2)+\varepsilon. \tag{42}$$

Analogously, for each $\varepsilon' > 0$ and for each $\Psi_2 \in \mathbb{L}_2$ there is a $\Psi_{\varepsilon'} \in \mathbb{L}_1$ s.t.

$$\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_2,\Psi_{\varepsilon'}) < \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_2,\mathbb{L}_1)+\varepsilon'. \tag{43}$$

Let us fix $\varepsilon, \varepsilon' > 0$. Then let $\Psi_\varepsilon \in \mathbb{L}_2$ be the formula realizing Equation (42), with respect to Ψ, and let $\tilde{\Psi_{\varepsilon'}}$ be the formula in $\mathbb{L}_1$ realizing Equation (42), with respect to this Ψ_ε. Therefore, we have

$$\begin{array}{lll}
 & |\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}_1) - \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}_2)| & \\
= & \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}_1) - \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}_2) & \\
< & \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}_1) - \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\Psi_\varepsilon)+\varepsilon & \text{(by Equation (42))} \\
= & \inf_{\Psi_1\in\mathbb{L}_1} \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\Psi_1) - \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\Psi_\varepsilon)+\varepsilon & \text{(by Definition 32)} \\
< & \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\tilde{\Psi_{\varepsilon'}}) - \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\Psi_\varepsilon)+\varepsilon & \\
\leq & \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\Psi_\varepsilon) + \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_\varepsilon,\tilde{\Psi_{\varepsilon'}}) - \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\Psi_\varepsilon)+\varepsilon & \text{(by triangle inequality)} \\
= & \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_\varepsilon,\tilde{\Psi_{\varepsilon'}})+\varepsilon & \\
< & \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_\varepsilon,\mathbb{L}_1)+\varepsilon'+\varepsilon & \text{(by Equation (43))} \\
\leq & \sup_{\Psi_2\in\mathbb{L}_2} \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi_2,\mathbb{L}_1)+\varepsilon'+\varepsilon & \\
\leq & \mathbf{H}(\mathscr{D}^{\mathrm{d}}_{\mathbb{L}})(\mathbb{L}_1,\mathbb{L}_2)+\varepsilon'+\varepsilon & \text{(by Equation (38)).}
\end{array}$$

Summarizing, we have obtained that

$$|\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}_1) - \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}_2)| < \mathbf{H}(\mathscr{D}^{\mathrm{d}}_{\mathbb{L}})(\mathbb{L}_1,\mathbb{L}_2)+\varepsilon'+\varepsilon$$

and since this inequality holds for each ε and ε', we can conclude that Equation (41) holds.

Equation (39) and Equation (40) taken together prove the thesis. □

To define the real-valued semantics of $\mathbb{L}^{\mathrm{d}}$ we exploit the distance $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}$. Informally, to quantify how much the formula Ψ is satisfied by process s we evaluate first how far Ψ is from being satisfied by s. This corresponds to the minimal distance between Ψ and a formula satisfied by s, namely to $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}(s))$. Then we simply notice that, as our distances are all 1-bounded, being $\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}(s))$ far from s is equivalent to be $1-\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}(s))$ close to it. Thus we assign to Ψ the real value $1-\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}(s))$ in s.

Definition 33 (Real-valued semantics of $\mathbb{L}^{\mathrm{d}}$)**.** We define the *real-valued semantics* of $\mathbb{L}^{\mathrm{d}}$ as the function $[\_](\_)\colon \mathbb{L}^{\mathrm{d}} \times \mathbf{S} \to [0,1]$ defined for all $\Psi \in \mathbb{L}^{\mathrm{d}}$ and $s \in \mathbf{S}$ as $[\Psi](s) = 1-\mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}(s))$.

We can restate our characterization theorem (Theorem 3) as a probabilistic $\mathbb{L}^{\mathrm{d}}$-model checking problem.

Theorem 12 (Characterization of strong trace metric II)**.** *For all $s,t \in \mathbf{S}$ we have*

$$\mathbf{d}_T(s,t) = \sup_{\Psi\in\mathbb{L}^{\mathrm{d}}} \mid [\Psi](s) - [\Psi](t) \mid .$$

Proof. From Theorem 3 we have $\mathbf{d}_T(s,t) = \mathscr{D}_{\mathbb{L}}(s,t)$. Hence the thesis is equivalent to prove

$$\mathscr{D}_{\mathbb{L}}(s,t) = \sup_{\Psi\in\mathbb{L}^{\mathrm{d}}} \mid [\Psi](s) - [\Psi](t) \mid .$$

We have

$$\begin{array}{rcll} \mathscr{D}_{\mathbb{L}}(s,t) & = & \mathbf{H}(\mathscr{D}^{\mathrm{d}}_{\mathbb{L}})(\mathbb{L}(s),\mathbb{L}(t)) & \text{(by Definition 28)} \\ & = & \sup_{\Psi\in\mathbb{L}^{\mathrm{d}}} \mid \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}(s)) - \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}(t)) \mid & \text{(by Proposition 15)} \\ & = & \sup_{\Psi\in\mathbb{L}^{\mathrm{d}}} \mid \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}(s)) - \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}(t)) + 1 - 1 \mid & \\ & = & \sup_{\Psi\in\mathbb{L}^{\mathrm{d}}} \mid 1 - \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}(t)) - \big(1 - \mathscr{D}^{\mathrm{d}}_{\mathbb{L}}(\Psi,\mathbb{L}(s))\big) \mid & \\ & = & \sup_{\Psi\in\mathbb{L}^{\mathrm{d}}} \mid [\Psi](t) - [\Psi](s) \mid & \text{(by Definition33).} \end{array}$$

□

8 Concluding remarks

We have provided a logical characterization of the strong and weak variants of trace metric on finite processes in the PTS model. Our results are based on the definition of a *distance* over the two-sorted boolean logics $\mathbb{L}$ and $\mathbb{L}_{\mathrm{w}}$, which we have proved to characterize resp. strong and weak probabilistic trace equivalence by exploiting the notion of *mimicking formula* of a resolution.

Our distance is a 1-bounded pseudometric that quantifies the syntactic disparities of the formulae and we have proved that the trace metric corresponds to the distance between the sets of formulae satisfied by the two processes. This approach, already successfully applied in [8] to the characterization of the bisimilarity metric, is not standard. Logical characterizations of the trace metrics have been obtained in terms of the probabilistic L-model checking problem, where L is the class of logical properties of interest, [1, 3, 9]. However we have proved that our approach can be exploited to regain classical one: by means of our distance between formulae we have defined a real-valued semantics for $\mathbb{L}$, namely a probabilistic model checking of a formula in a process, and then we have proved that the trace metric constitutes the least upper bound to the error that can be observed in the verification of an $\mathbb{L}$ formula.

Another interesting feature of our approach is its generality, since it can be easily applied to some variants of the trace equivalence and trace metric. In [4, 26] the authors distinguish between resolutions obtained via deterministic schedulers and the ones obtained via randomized schedulers. The only difference between the two classes is in the evaluation of the probability weights: in deterministic resolutions, which are the ones we have considered in this paper, each possible resolution of nondeterminism is considered singularly and thus the target probability distributions of their transitions are the same as in the considered process. In randomized resolutions, internal nondeterminism is solved by assigning a probability weight to each choice and thus the target distributions are obtained from the convex combination of the target distributions of the considered process. Since the definition of the mimicking formulae depends solely on the values of the probability weights in the resolutions and not on how these weights are evaluated, our characterization can be applied also to the case of trace equivalences and metrics defined in terms of randomized resolutions.

As a first step in the future development of our work, we aim to extend our results to the trace equivalence defined in [4] which, differently from the equivalence of [26] considered in this paper, is compositional wrt. the parallel composition operator. Roughly speaking, in [4] for each given trace it is checked whether the resolutions of two processes assign the same probability it, whereas in [26] for a chosen resolution of the first process we check whether there is a resolution for the second process that assigns the same probability to all traces. Furthermore, no trace metric has been defined yet for the equivalence in [4]. Our idea is then firstly to define such a trace metric and secondly to simplify the logic $\mathbb{L}$ by substituting the trace distribution formulae with a simple test on the execution probability of a trace, with an operator similar to the probabilistic operator in [25]. By applying our approach to the new logic we will obtain the characterization of the trace equivalence and metric.

Then, we will study metrics and logical characterizations for the testing equivalences defined in [4].

Finally, in [3] a sequence of Kantorovich bisimilarity-like metrics converging to the trace metric on MCs is provided. Hence we aim to combine our characterization results in [8] with the ones in this paper in order to see if a similar result of convergence can be obtained also with our technique on PTSs.

References

[1] Luca de Alfaro, Marco Faella & Mariëlle Stoelinga (2009): *Linear and Branching System Metrics*. IEEE Trans. Software Eng. 35(2), pp. 258–273, doi:10.1109/TSE.2008.106.

[2] Luca de Alfaro, Rupak Majumdar, Vishwanath Raman & Mariëlle Stoelinga (2008): *Game Refinement Relations and Metrics*. Logical Methods in Computer Science 4(3).

[3] Giorgio Bacci, Giovanni Bacci, Kim G. Larsen & Radu Mardare (2015): *Converging from Branching to Linear Metrics on Markov Chains*. In: Proceedings of ICTAC 2015, pp. 349–367, doi:10.1007/978-3-319-25150-9_21.

[4] Marco Bernardo, Rocco De Nicola & Michele Loreti (2014): *Revisiting Trace and Testing Equivalences for Nondeterministic and Probabilistic Processes*. Logical Methods in Computer Science 10(1), doi:10.2168/LMCS-10(1 : 16)2014.

[5] Franck van Breugel (2005): *A Behavioural Pseudometric for Metric Labelled Transition Systems*. In: Proceedings of CONCUR 2005, pp. 141–155, doi:10.1007/11539452_14.

[6] Franck van Breugel & James Worrell (2001): *Towards Quantitative Verification of Probabilistic Transition Systems*. In: Proceedings of ICALP, pp. 421–432, doi:10.1007/3-540-48224-5_35.

[7] Franck van Breugel & James Worrell (2005): *A behavioural pseudometric for probabilistic transition systems*. Theor. Comput. Sci. 331(1), pp. 115–142.

[8] Valentina Castiglioni, Daniel Gebler & Simone Tini (2016): *Logical Characterization of Bisimulation Metrics*. In: Proceedings of QAPL 2016, Electronic Proceedings in Theoretical Computer Science, doi:10.4204/EPTCS.227.4.

[9] Przemyslaw Daca, Thomas A. Henzinger, Jan Kretínský & Tatjana Petrov (2016): *Linear Distances between Markov Chains*. In: Proceedings of CONCUR 2016, pp. 20:1–20:15, doi:10.4230/LIPIcs.CONCUR.2016.20.

[10] Yuxin Deng, Tom Chothia, Catuscia Palamidessi & Jun Pang (2006): *Metrics for Action-labelled Quantitative Transition Systems*. Electronic Notes in Theoretical Computer Science 153(2), pp. 79–96, doi:10.1016/j.entcs.2005.10.033.

[11] Yuxin Deng & Wenjie Du (2011): *Logical, Metric, and Algorithmic Characterisations of Probabilistic Bisimulation*. CoRR abs/1103.4577.

[12] Josee Desharnais, Vineet Gupta, Radha Jagadeesan & Prakash Panangaden (2004): *Metrics for labelled Markov processes*. Theoretical Computer Science 318(3), pp. 323–354, doi:10.1016/j.tcs.2003.09.013.

[13] Josée Desharnais, Radha Jagadeesan, Vineet Gupta & Prakash Panangaden (2002): *The Metric Analogue of Weak Bisimulation for Probabilistic Processes*. In: Proc. LICS 2002, pp. 413–422.

[14] Wenjie Du, Yuxin Deng & Daniel Gebler (2016): *Behavioural Pseudometrics for Nondeterministic Probabilistic Systems*. In: Proceedings of SETTA 2016, pp. 67–84, doi:10.1007/978-3-319-47677-3_5.

[15] D. Gebler, K. G. Larsen & S. Tini (2016): *Compositional bisimulation metric reasoning with Probabilistic Process Calculi*. Logical Methods in Computer Science 12(4).

[16] Daniel Gebler & Simone Tini (2015): *SOS Specifications of Probabilistic Systems by Uniformly Continuous Operators*. In: Proc. CONCUR 2015, pp. 155–168.

[17] Alessandro Giacalone, Chi-Chang Jou & Scott A. Smolka (1990): *Algebraic Reasoning for Probabilistic Concurrent Systems*. In: Proc. IFIP Work, Conf. on Programming, Concepts and Methods, pp. 443–458.

[18] Hans Hansson & Bengt Jonsson (1994): *A logic for reasoning about time and reliability*. *Formal Aspects of Computing* 6(5), pp. 512–535, doi:10.1007/BF01211866.

[19] Matthew Hennessy & Robin Milner (1985): *Algebraic laws for nondeterminism and concurrency*. *J. Assoc. Comput. Mach.* 32, pp. 137–161.

[20] Holger Hermanns, Augusto Parma, Roberto Segala, Björn Wachter & Lijun Zhang (2011): *Probabilistic Logical Characterization*. *Information and Computation* 209(2), pp. 154–172, doi:10.1016/j.ic.2010.11.024.

[21] Leonid V. Kantorovich (1942): *On the Transfer of Masses*. Original article in Russian, translation in Management Science, $5:1-4(1959)$.

[22] Robert M. Keller (1976): *Formal Verification of Parallel Programs*. *Commun. ACM* 19(7), pp. 371–384, doi:10.1145/360248.360251.

[23] Marta Z. Kwiatkowska & Gethin Norman (1996): *Probabilistic Metric Semantics for a Simple Language with Recursion*. In: *Proc. MFCS'96*, pp. 419–430.

[24] Kim G. Larsen, Radu Mardare & Prakash Panangaden (2012): *Taking It to the Limit: Approximate Reasoning for Markov Processes*. In: *Proc. MFCS 2012*, pp. 681–692.

[25] Augusto Parma & Roberto Segala (2007): *Logical Characterizations of Bisimulations for Discrete Probabilistic Systems*. In: *Proceedings of FoSSaCS 2007*, pp. 287–301, doi:10.1007/978-3-540-71389-0_21.

[26] Roberto Segala (1995): *A Compositional Trace-Based Semantics for Probabilistic Automata*. In: *Proceedings of CONCUR '95*, pp. 234–248, doi:10.1007/3-540-60218-6_17.

[27] Roberto Segala (1995): *Modeling and Verification of Randomized Distributed Real-Time Systems*. Ph.D. thesis, MIT.

[28] Lin Song, Yuxin Deng & Xiaojuan Cai (2007): *Towards Automatic Measurement of Probabilistic Processes*. In: *Proceedings of QSIC 2007*, pp. 50–59, doi:10.1109/QSIC.2007.65.

[29] William J. Stewart (1994): *Introduction to the Numerical Solution of Markov Chains*. Princeton University Press.