On the Tradeoff between Compositionality and Exactness in Weak Bisimilarity for Integrated-Time Markovian Process Calculi

Marco Bernardo

Dipartimento di Scienze di Base e Fondamenti – Università di Urbino – Italy

Abstract

Integrated-time Markovian process calculi rely on actions whose durations are quantified by exponentially distributed random variables. The Markovian bisimulation equivalences defined so far for these calculi treat exponentially timed internal actions like all the other actions, because each such action has a nonzero duration and hence can be observed if it is executed between a pair of exponentially timed noninternal actions. However, no difference may be noted, at stationary state, between a sequence of exponentially timed internal actions and a single exponentially timed internal action, if their expected durations and execution probabilities coincide, a fact exploited in Hillston's weak isomorphism. We show that Milner's approach can be adapted on the basis of this fact, so to derive a weak bisimulation equivalence for integrated-time Markovian process calculi, up to a tradeoff between compositionality and exactness inherent to the Markovian setting. The resulting weak Markovian bisimulation equivalence induces a pseudo-aggregation that is exact at stationary state for all the considered processes, but turns out to be a congruence only over sequential processes. To achieve compositionality over concurrent processes, we need to enhance the abstraction capability of the equivalence in the presence of interleaved computations. However, the corresponding pseudo-aggregation turns out to be exact at stationary state only for a subset of concurrent processes. In addition to this tradeoff, we present, for the first equivalence, a sound and complete axiomatization over sequential processes, which is instrumental to characterize pseudo-aggregations, and a polynomial-time equivalence-checking algorithm, which can be exploited for the compositional minimization of concurrent processes.

Keywords: stochastic process algebra, weak bisimulation equivalence, compositionality, continuous-time Markov chains, pseudo-aggregations, exactness

1. Introduction

Quantitative models based on continuous-time Markov chains (see, e.g., [34]) like stochastic Petri nets (see, e.g., [2]) and stochastic process algebras (see, e.g., [22, 20]) have been deeply investigated and successfully used in the last decades to predict the performance of computer, communication, and software systems. From a conceptual viewpoint, we can distinguish between *integrated-time* and *orthogonal-time* Markovian models [9]. In the former, which are more natural for modeling purposes, the passage of time is associated with the execution of activities, i.e., activities are considered *durational*. In the latter, which are more elegant on the theoretical side, the passage of time is separate from the execution of activities, i.e., activities are *durationless* and hence time passing has to be represented explicitly.

Several Markovian behavioral equivalences (see [3] and the references therein) have been proposed in the literature for relating and manipulating system models with an underlying continuous-time Markov chain (CTMC) semantics. These equivalences are extensions of the traditional approaches to the definition of behavioral equivalences, and take into account time passing described by means of exponential distributions. A feature shared by relations like Markovian bisimilarity, Markovian testing equivalence, and Markovian trace equivalence is that of being *strong*, in the sense that they treat internal activities – which cannot be seen by an external observer – like the other activities. Only a few variants investigated in [20, 31, 25, 12] are able to abstract from internal activities and/or purely probabilistic branchings.

Preprint submitted to Theoretical Computer Science *October 25, 2014*

The useful capability of abstracting from internal actions can be easily achieved in the orthogonal-time setting, because in that case activities are immediate (i.e., take no time) and hence well-known techniques developed for nondeterministic processes can be employed to get rid of these activities when they are internal. Let us denote by τ the invisible or silent action. In the nondeterministic setting, a process that can perform action a followed by action τ and action b and then terminates – written $a \,.\, \tau \,.\, b \,.\, \underline{0}$ – is weakly equivalent to a process that can perform action a followed by action b and then terminates – written $a \,.\, b \,.\, \underline{0}$. The situation is more complicated in the integrated-time setting. Since actions have exponentially distributed durations – uniquely identified by positive real numbers called rates – it is not necessarily the case that simplifications like the one above can be made.

For instance, if action a has rate λ, action b has rate μ, and action τ has rate γ, the two resulting integrated-time Markovian processes $<a, \lambda>.<\tau, \gamma>.<b, \mu>.\underline{0}$ and $<a, \lambda>.<b, \mu>.\underline{0}$ are not weakly equivalent. In fact, recalling that the expected duration of an action coincides with the reciprocal of the rate of the action, the former process has a maximal computation whose expected duration is $\frac{1}{\lambda} + \frac{1}{\gamma} + \frac{1}{\mu}$, whereas the latter process has a maximal computation whose expected duration is $\frac{1}{\lambda} + \frac{1}{\mu}$. From another viewpoint, in the former case an external observer would see an a-action for an amount of time t_λ and a b-action for an amount of time t_μ, with a delay t_γ in between, while in the latter case the external observer would not see any delay between the termination of the execution of a and the beginning of the execution of b. Therefore, in a Markovian setting, a τ-action executed between a pair of non-τ-actions cannot be abstracted away, because it has a nonzero duration and hence can be, from a timing viewpoint, observed.

Hillston's weak isomorphism [22] indicates that we should not be too pessimistic. As a different example, take a process that, between actions a and b, can perform *two* τ-actions with rates γ_1 and γ_2, respectively: $<a, \lambda>.<\tau, \gamma_1>.<\tau, \gamma_2>.<b, \mu>.\underline{0}$. In this case, an observer may not be able to distinguish between the execution of the two τ-actions above and the execution of *a single τ-action whose expected duration is the sum of the expected durations of the two original τ-actions*, i.e., $\frac{1}{\gamma_1} + \frac{1}{\gamma_2} = \frac{\gamma_1 + \gamma_2}{\gamma_1 \cdot \gamma_2}$. In other words, the process may be viewed as being *weakly equivalent* to $<a, \lambda>.<\tau, \frac{\gamma_1 \cdot \gamma_2}{\gamma_1 + \gamma_2}>.<b, \mu>.\underline{0}$.

The two processes above are certainly weakly equivalent from a functional standpoint. However, since the sum of the two exponential random variables quantifying the durations of the two original τ-actions has been *approximated* with a single mean-preserving exponential random variable, it is not necessarily the case that the two processes have the same performance characteristics. This would be true if the equivalence induced a *pseudo*-aggregation of the underlying CTMC that is exact, i.e., such that the transient/stationary probability of being in a macrostate of the aggregated stochastic process – which is *assumed* to be a CTMC – is the sum of the transient/stationary probabilities of being in one of the constituent microstates of the original CTMC. This is the case with Markovian bisimilarity, which is in agreement with the well-known exact CTMC-level aggregation called ordinary lumpability [22, 16], and Markovian testing and trace equivalences, which are consistent with a coarser exact CTMC-level aggregation called T-lumpability [8, 33].

In this paper, we show that the construction used in [27] to derive a weak bisimulation equivalence for nondeterministic process calculi can be extended to integrated-time Markovian process calculi. The resulting equivalence is weak in the sense that it is capable of *abstracting from the number and the order of consecutive exponentially timed τ-actions* in a computation. It reduces any such sequence to a *single* exponentially timed τ-action *preserving both the expected duration and the execution probability of the original action sequence*. From a stochastic viewpoint, this reduction amounts to replacing hypoexponentially distributed durations with exponentially distributed durations having the same expected value. As a consequence, processes related by the resulting equivalence will not possess the same transient performance measures, unless they refer to properties expressed as the *mean time to certain events*. However, those processes may possess the same stationary reward-based performance measures, as the pseudo-aggregation induced by the considered equivalence on the CTMC underlying each process may be *exact at stationary state*.

Defining a weak Markovian bisimilarity that works as outlined above causes a *tradeoff between semantical compositionality and pseudo-aggregation exactness* to emerge, which is inherent to the Markovian setting. For this reason, we divide the presentation of our results into two parts.

Firstly, we extend the construction of [27] in the simplest possible way, so that the only sequences of exponentially timed internal transitions that are reduced are those that traverse *states enabling only* *exponentially timed internal actions*. The resulting weak Markovian bisimulation equivalence induces a pseudo-aggregation – called W-lumpability – that is exact at stationary state for *all* the considered processes, thus ensuring full preservation of stationary reward-based performance measures. However, the equivalence is a congruence *only over sequential processes*, a fact that limits its usefulness for state space minimization purposes when there are several processes composed in parallel.

Secondly, we retrieve compositionality over concurrent processes by enhancing the abstraction capability of the equivalence in the presence of interleaved computations. Given a sequential process, if it originates a sequence of exponentially timed internal transitions that traverse *local states* enabling only exponentially timed internal actions, the basic idea is to apply the reduction also when that process is composed in parallel with other processes, and hence the sequence may traverse *global states that enable observable actions too*. The resulting generalized weak Markovian bisimulation equivalence is shown to be a congruence over *all* the considered processes. However, the induced generalized pseudo-aggregation – called GW-lumpability – turns out to be exact at stationary state *only for a subset* of the considered processes, which are those with fully independent or fully synchronized sequential components, and those in which only certain synchronizations take place before the sequences to be reduced.

For the first equivalence, we also exhibit a sound and complete axiomatization over sequential processes, which has been instrumental to characterize W-lumpability and GW-lumpability. Moreover, we show that it is decidable in polynomial time for finite-state processes having no cycles of exponentially timed internal transitions, and exemplify how to exploit the corresponding equivalence-checking algorithm for the compositional minimization of concurrent processes according to the generalized equivalence.

This paper, which is an extended and revised version of [10, 11], is organized as follows. In Sect. 2, we introduce an integrated-time Markovian process calculus and recall Markovian bisimilarity. In Sect. 3, we develop a weak variant of Markovian bisimilarity that achieves full exactness at stationary state, but only a limited form of compositionality. In Sect. 4, we enhance the abstraction capability so to obtain full compositionality, at the price of losing exactness for a subset of the considered processes. In Sect. 5, we provide an algorithm for deciding the first weak Markovian bisimilarity, then we illustrate its use for compositional state space minimization with respect to the second weak Markovian bisimilarity. In Sect. 6, we discuss related work, in particular Hillston's weak isomorphism. Finally, in Sect. 7 we provide some concluding remarks. For the sake of readability, all proofs are collected in an appendix.

2. Integrated-Time Markovian Process Calculi and Markovian Bisimilarity

In order to study properties like compositionality and axiomatizability of weak Markovian bisimilarity, it is convenient to define a Markovian process calculus (MPC for short). In this calculus, we firstly include the operators that are necessary to generate all the action-labeled CTMCs: the inactive process, exponentially timed action prefix, alternative composition, and recursion. These operators, some referred to as *dynamic*, results in sequential components. In addition, we include *static* operators such as parallel composition, because we are also interested in concurrent systems, and hiding, because the behavioral equivalence we are going to propose is weak and hence we need a way to make actions invisible.

In the integrated-time setting, an action is represented as a pair $<a, \lambda>$. The first element, a, is the name of the action, which is τ in the case that the action is internal, otherwise it belongs to a set $Name_{\mathrm{v}}$ of visible action names. The second element, $\lambda \in \mathbb{R}_{>0}$, is the rate of the exponentially distributed random variable RV quantifying the duration of the action, i.e., $\Pr\{RV \leq t\} = 1 - \mathrm{e}^{-\lambda \cdot t}$ for $t \in \mathbb{R}_{>0}$, with the expected duration of the action being equal to $1/\lambda$.

If several exponentially timed actions are simultaneously enabled, the action that is executed is the one sampling the least duration. This mechanism – called *race policy* – implies that the sojourn time associated with a process term P is the minimum of the random variables quantifying the durations of the exponentially timed actions enabled by P. Such a minimum turns out to be exponentially distributed, with rate equal to the sum of the rates of the actions enabled by P. Therefore, the expected sojourn time associated with P is the reciprocal of the sum of the rates of the actions it enables. The probability of executing one of those actions is given by the action rate divided by the sum of the rates of all the considered actions.

$$
\begin{array}{llll}
(\textsc{Pre}) & \dfrac{}{<a,\lambda>.P \xrightarrow{a,\lambda} P} & (\textsc{Rec}) & \dfrac{P\{\mathrm{rec}\, X : P \hookrightarrow X\} \xrightarrow{a,\lambda} P'}{\mathrm{rec}\, X : P \xrightarrow{a,\lambda} P'} \\[2ex]
(\textsc{Alt}_1) & \dfrac{P_1 \xrightarrow{a,\lambda} P'}{P_1 + P_2 \xrightarrow{a,\lambda} P'} & (\textsc{Alt}_2) & \dfrac{P_2 \xrightarrow{a,\lambda} P'}{P_1 + P_2 \xrightarrow{a,\lambda} P'} \\[2ex]
(\textsc{Par}_1) & \dfrac{P_1 \xrightarrow{a,\lambda} P_1' \quad a \notin S}{P_1 \,\|_S\, P_2 \xrightarrow{a,\lambda} P_1' \,\|_S\, P_2} & (\textsc{Par}_2) & \dfrac{P_2 \xrightarrow{a,\lambda} P_2' \quad a \notin S}{P_1 \,\|_S\, P_2 \xrightarrow{a,\lambda} P_1 \,\|_S\, P_2'}
\end{array}
$$

$$
(\textsc{Syn}) \quad \frac{P_1 \xrightarrow{a,\lambda_1} P_1' \quad P_2 \xrightarrow{a,\lambda_2} P_2' \quad a \in S}{P_1 \,\|_S\, P_2 \xrightarrow{a,\lambda_1 \otimes \lambda_2} P_1' \,\|_S\, P_2'}
$$

$$
\begin{array}{llll}
(\textsc{Hid}_1) & \dfrac{P \xrightarrow{a,\lambda} P' \quad a \notin H}{P/H \xrightarrow{a,\lambda} P'/H} & (\textsc{Hid}_2) & \dfrac{P \xrightarrow{a,\lambda} P' \quad a \in H}{P/H \xrightarrow{\tau,\lambda} P'/H}
\end{array}
$$

Table 1: Structural operational semantic rules for process terms in $\mathbb{P}$

Definition 2.1. Let $Act = Name \times \mathbb{R}_{>0}$ be a set of actions, where $Name = Name_{\mathrm{v}} \cup \{\tau\}$ is a set of action names – ranged over by a, b – and $\mathbb{R}_{>0}$ is a set of action rates – ranged over by λ, μ, γ. Let Var be a set of process variables – ranged over by X, Y. The process language $\mathcal{PL}$ is generated by the following syntax:

P ::=	C	sequential component
|	$P \,\|_S\, P$	parallel composition
|	P/H	hiding
C ::=	$\underline{0}$	inactive process
|	$<a,\lambda>.C$	exponentially timed action prefix
|	$C + C$	alternative composition
|	X	process variable
|	$\mathrm{rec}\, X : C$	recursion

where $S, H \subseteq Name_{\mathrm{v}}$. We denote by $\mathbb{P}$ the set of closed and guarded process terms of $\mathcal{PL}$ – ranged over by P, Q – with $\mathbb{P}_{\mathrm{seq}}$ being its subset generated by the C-production. ■

In order to distinguish between process terms like $<a,\lambda>.\underline{0} + <a,\lambda>.\underline{0}$ and $<a,\lambda>.\underline{0}$, the semantic model $[\![P]\!]$ for a process term $P \in \mathbb{P}$ is a labeled *multi*transition system as it takes into account the multiplicity of each transition, intended as the number of different proofs for the transition derivation. The multitransition relation of $[\![P]\!]$ is contained in the smallest multiset of elements of $\mathbb{P} \times Act \times \mathbb{P}$ that satisfies the operational semantic rules in Table 1 – where $\{\_ \hookrightarrow \_\}$ denotes syntactical replacement – *and* keeps track of all the possible ways of deriving each of its transitions. With regard to rule Syn, we assume that the duration of an action deriving from the synchronization of two exponentially timed actions is exponentially distributed, with a rate obtained by applying (like, e.g., in [21]) some commutative and associative operation denoted by $\otimes$ to the rates of the two original actions.

The notion of bisimilarity for MPC is based on the comparison of exit rates [22, 21]. The *exit rate* of a process term $P \in \mathbb{P}$ with respect to action name $a \in Name$ and destination $D \subseteq \mathbb{P}$ is the rate at which P can execute actions of name a that lead to D:

$$
\boxed{rate(P, a, D) \;=\; \sum \{\!|\, \lambda \in \mathbb{R}_{>0} \mid \exists P' \in D.\, P \xrightarrow{a,\lambda} P' \,|\!\}}
$$

where $\{\!|$ and $|\!\}$ are multiset delimiters and the summation is taken to be zero if the multiset is empty. By summing up the rates of all the actions of P, we obtain the *total exit rate* of P:

$$\boxed{rate_{\mathrm{t}}(P) \;=\; \sum_{a \in Name} rate(P, a, \mathbb{P})}$$

which is the reciprocal of the expected sojourn time associated with P.

Definition 2.2. An equivalence relation $\mathcal{B}$ over $\mathbb{P}$ is a *Markovian bisimulation* iff, whenever $(P_1, P_2) \in \mathcal{B}$, then for all action names $a \in Name$ and equivalence classes $D \in \mathbb{P}/\mathcal{B}$:

$$rate(P_1, a, D) \;=\; rate(P_2, a, D)$$

We call *Markovian bisimilarity*, denoted by $\sim_{\mathrm{MB}}$, the largest Markovian bisimulation. ■

The relation $\sim_{\mathrm{MB}}$ possesses the following properties:

- $\sim_{\mathrm{MB}}$ is a congruence with respect to all the operators of MPC as well as recursion [22, 21, 15].
- $\sim_{\mathrm{MB}}$ has a sound and complete axiomatization [22, 21]. Its basic laws for dynamic operators are:

$$\boxed{\begin{array}{lrcl}
(\mathcal{A}_{\mathrm{MB},1}) & P_1 + P_2 & = & P_2 + P_1 \\
(\mathcal{A}_{\mathrm{MB},2}) & (P_1 + P_2) + P_3 & = & P_1 + (P_2 + P_3) \\
(\mathcal{A}_{\mathrm{MB},3}) & P + \underline{0} & = & P \\
(\mathcal{A}_{\mathrm{MB},4}) & <a, \lambda_1>.P + <a, \lambda_2>.P & = & <a, \lambda_1 + \lambda_2>.P
\end{array}}$$

  where the last one encodes the race policy and hence replaces the idempotency law $P + P = P$ valid for nondeterministic processes. The other laws are the expansion law for parallel composition and the distribution laws for hiding or, alternatively, the unfolding laws for recursion.
- $\sim_{\mathrm{MB}}$ induces a well-known CTMC-level aggregation called ordinary lumpability, which is exact both at stationary state and at transient state [22, 16].
- $\sim_{\mathrm{MB}}$ can be decided in polynomial time for all finite-state processes [35].

3. Abstracting from Internal Actions: Full Exactness, Limited Compositionality

We now weaken the distinguishing power of $\sim_{\mathrm{MB}}$ in order to abstract from exponentially timed τ-transitions. As noted in Sect. 1, while it is not possible to get rid of an individual exponentially timed τ-action executed between a pair of exponentially timed non-τ-actions, the execution of a sequence of exponentially timed τ-actions may be indistinguishable, at stationary state, from the execution of a single exponentially timed τ-action having the same expected duration and execution probability as the sequence. Based on this consideration and the construction of [27], in Sect. 3.1 we define a weak variant of $\sim_{\mathrm{MB}}$ over MPC, which in Sect. 3.2 we prove to be a congruence except for parallel composition. In Sect. 3.3, we exhibit a sound and complete axiomatization over nonrecursive sequential process terms, which is instrumental to demonstrate in Sect. 3.4 that the equivalence induces a pseudo-aggregation that is exact at stationary state for all processes. Finally, in Sect. 3.5 we discuss coarser variants of the equivalence.

3.1. Definition of $\approx_{\mathrm{MB}}$

We say that $P \in \mathbb{P}$ is *stable* if $P \overset{\tau,\lambda}{\not\longrightarrow} P'$ for all λ and P', otherwise we say that P is *unstable*. In the latter case, we say that P is *fully unstable* iff, whenever $P \xrightarrow{a,\lambda} P'$, then $a = \tau$. We let $\mathbb{P} = \mathbb{P}_{\mathrm{nfu}} \cup \mathbb{P}_{\mathrm{fu}}$, where $\mathbb{P}_{\mathrm{nfu}}$ and $\mathbb{P}_{\mathrm{fu}}$ are the sets of process terms that are not fully unstable and fully unstable, respectively. From now on, we concentrate on sequences of exponentially timed τ-actions labeling computations that traverse fully unstable states, as they are the most natural candidates for abstraction purposes.

Definition 3.1. Let $n \in \mathbb{N}_{\geq 1}$ and $P_1, P_2, \ldots, P_{n+1} \in \mathbb{P}$. A computation c of length n from P_1 to P_{n+1} having the form $P_1 \xrightarrow{\tau,\lambda_1} P_2 \xrightarrow{\tau,\lambda_2} \ldots \xrightarrow{\tau,\lambda_n} P_{n+1}$ is *reducible* iff $P_i \in \mathbb{P}_{\rm fu}$ for all $i = 1, \ldots, n$. ■

If reducible, the computation c above can be reduced to a single exponentially timed τ-transition tr that we consider equivalent to c if the rate of tr subsumes the execution probability of c (product of the execution probabilities of the transitions of c) and the expected duration of c (sum of the expected sojourn times in the states traversed by c). The rate of tr is obtained from the positive real value below:

$$\boxed{\mathit{probtime}(c) \; = \; \left(\prod_{i=1}^{n} \frac{\lambda_i}{\mathit{rate}_{\rm t}(P_i)}\right) \cdot \left(\sum_{i=1}^{n} \frac{1}{\mathit{rate}_{\rm t}(P_i)}\right)}$$

by leaving its first factor unchanged – which is the execution probability of c – and taking the reciprocal of the second factor – which is the expected duration of c, with $P_i \in \mathbb{P}_{\rm fu}$ implying $\mathit{rate}_{\rm t}(P_i) = \mathit{rate}(P_i, \tau, \mathbb{P})$. In other words, $\mathit{probtime}(c)$ is the expected duration of c weighted by the execution probability of c itself. For example, if we consider the reducible computation c of $<\tau, \mu>.<\tau, \gamma>.\underline{0}$ made out of both exponentially timed τ-transitions, then $\mathit{probtime}(c) = (\frac{\mu}{\mu} \cdot \frac{\gamma}{\gamma}) \cdot (\frac{1}{\mu} + \frac{1}{\gamma}) = 1 \cdot \frac{\mu+\gamma}{\mu\cdot\gamma}$ and hence the equivalent rate is $\frac{\mu\cdot\gamma}{\mu+\gamma}$.

As can be noted, we consider only reducible computations of finite length. This will be enough to distinguish between fully unstable process terms that must be told apart. In fact, assuming $\lambda_1 \neq \lambda_2$, it makes sense to discriminate between $<\tau, \lambda_1>.P$ and $<\tau, \lambda_2>.P$ if P can reach a non-fully-unstable process term. By contrast, an external observer cannot see any difference between two divergent process terms such as $\mathrm{rec}\, X : <\tau, \lambda_1>.X$ and $\mathrm{rec}\, X : <\tau, \lambda_2>.X$ because they do not reach any non-fully-unstable process term.

We are now ready to define a weak variant of $\sim_{\rm MB}$ such that (i) processes in $\mathbb{P}_{\rm nfu}$ are dealt with as in $\sim_{\rm MB}$ and (ii) the length of reducible computations from processes in $\mathbb{P}_{\rm fu}$ to processes in $\mathbb{P}_{\rm nfu}$ is abstracted away while preserving their execution probability and expected duration. In the latter case, we need to lift measure *probtime* from individual reducible computations to multisets of reducible computations, which requires summing up the *probtime* measures of those computations whenever appropriate. More precisely, denoting by $\mathit{rcomp}(P, D, t)$ the multiset of reducible computations from $P \in \mathbb{P}_{\rm fu}$ to some P' in $D \subseteq \mathbb{P}$ whose expected duration is $t \in \mathbb{R}_{>0}$, we consider the following t-indexed multiset of sums of *probtime* measures:

$$\boxed{\mathit{pbtm}(P, D) \; = \bigcup_{t \in \mathbb{R}_{>0} \text{ s.t. } \mathit{rcomp}(P,D,t) \neq \emptyset} \{\!| \sum_{c \in \mathit{rcomp}(P,D,t)} \mathit{probtime}(c) \,|\!\}}$$

Notice that $\mathit{pbtm}(P, D)$ is not simply the multiset of the *probtime* measures of the various reducible computations from P to D. In that case, e.g., we would have $\mathit{pbtm}(<\tau, \lambda_1>.\underline{0} + <\tau, \lambda_2>.\underline{0}, \{\underline{0}\}) = \{\!| \frac{\lambda_1}{\lambda_1+\lambda_2} \cdot \frac{1}{\lambda_1+\lambda_2}, \frac{\lambda_2}{\lambda_1+\lambda_2} \cdot \frac{1}{\lambda_1+\lambda_2} |\!\}$ while $\mathit{pbtm}(<\tau, \lambda_1 + \lambda_2>.\underline{0}, \{\underline{0}\}) = \{\!| \frac{1}{\lambda_1+\lambda_2} |\!\}$, thus obtaining a behavioral equivalence that is not a conservative extension of $\sim_{\rm MB}$. As a consequence, in the definition of $\mathit{pbtm}(P, D)$, the various *probtime* measures are summed up over all reducible computations from P to D *having the same expected duration* t. A more radical option would be to sum up the *probtime* measures of all the computations from P to D *regardless of their expected durations*; this will be discussed in Sect. 3.5.

Definition 3.2. An equivalence relation $\mathcal{B}$ over $\mathbb{P}$ is a *weak Markovian bisimulation* iff, whenever $(P_1, P_2) \in \mathcal{B}$, then one of the following holds:

- $P_1, P_2 \in \mathbb{P}_{\rm nfu}$ and for all $a \in \mathit{Name}$ and equivalence classes $D \in \mathbb{P}/\mathcal{B}$:
$$\mathit{rate}(P_1, a, D) \; = \; \mathit{rate}(P_2, a, D)$$
- $P_1, P_2 \in \mathbb{P}_{\rm fu}$ and for all equivalence classes $D \in \mathbb{P}_{\rm nfu}/\mathcal{B}$:
$$\mathit{pbtm}(P_1, D) \; = \; \mathit{pbtm}(P_2, D)$$

We call *weak Markovian bisimilarity*, denoted by $\approx_{\rm MB}$, the largest weak Markovian bisimulation. ■

Relation $\approx_{\rm MB}$ cannot equate a fully unstable, divergent process term like $\mathrm{rec}\, X : <\tau, \lambda>.X$ to a fully unstable, non-divergent process term like $<\tau, \lambda>.\underline{0}$, as only the latter can reach a non-fully-unstable process term, which is $\underline{0}$. As observed in [20], this is important for compositionality purposes in a Markovian setting.

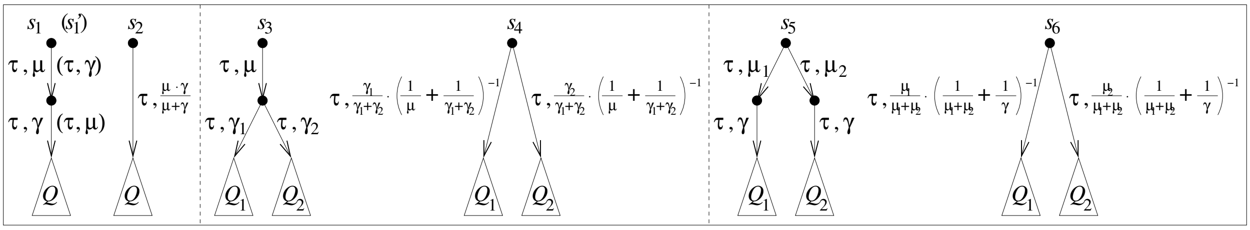

Figure 1: Additional identifications made by $\approx_{\rm MB}$ with respect to $\sim_{\rm MB}$ (see Exs. 3.3, 3.4, and 3.5)

We provide below a number of examples that should clarify the additional identifications made by $\approx_{\rm MB}$ with respect to $\sim_{\rm MB}$ and the role of the two factors of *probtime*. The identifications are illustrated in Fig. 1.

Example 3.3. Consider the following two process terms:

$$\begin{array}{l} \bar{P}_1 \;\equiv\; <\tau,\mu>.<\tau,\gamma>.Q \qquad (\text{or } \bar{P}'_1 \;\equiv\; <\tau,\gamma>.<\tau,\mu>.Q) \\ \bar{P}_2 \;\equiv\; <\tau,\frac{\mu\cdot\gamma}{\mu+\gamma}>.Q \end{array}$$

with $Q \in \mathbb{P}_{\rm nfu}$. As anticipated in Sect. 1, it turns out that $\bar{P}_1 \approx_{\rm MB} \bar{P}_2$ because:

$$pbtm(\bar{P}_1,[Q]_{\approx_{\rm MB}}) \;=\; \{\!|\, (1\cdot 1)\cdot(\tfrac{1}{\mu}+\tfrac{1}{\gamma}) \,|\!\} \;=\; \{\!|\, 1\cdot\tfrac{\mu+\gamma}{\mu\cdot\gamma} \,|\!\} \;=\; pbtm(\bar{P}_2,[Q]_{\approx_{\rm MB}})$$

where $[Q]_{\approx_{\rm MB}}$ is the equivalence class of Q with respect to $\approx_{\rm MB}$.
In general, for $l \in \mathbb{N}_{\ge 1}$ we have that $<\tau,\mu>.<\tau,\gamma_1>.\dots.<\tau,\gamma_l>.Q$ is weakly Markovian bisimilar to $<\tau,\left(\frac{1}{\mu}+\frac{1}{\gamma_1}+\dots+\frac{1}{\gamma_l}\right)^{-1}>.Q$. ■

Example 3.4. Consider the following two process terms:

$$\begin{array}{l} \bar{P}_3 \;\equiv\; <\tau,\mu>.(<\tau,\gamma_1>.Q_1 + <\tau,\gamma_2>.Q_2) \\ \bar{P}_4 \;\equiv\; <\tau,\frac{\gamma_1}{\gamma_1+\gamma_2}\cdot\left(\frac{1}{\mu}+\frac{1}{\gamma_1+\gamma_2}\right)^{-1}>.Q_1 + <\tau,\frac{\gamma_2}{\gamma_1+\gamma_2}\cdot\left(\frac{1}{\mu}+\frac{1}{\gamma_1+\gamma_2}\right)^{-1}>.Q_2 \end{array}$$

with $Q_1, Q_2 \in \mathbb{P}_{\rm nfu}$ and $Q_1 \not\approx_{\rm MB} Q_2$. Unlike action $<\tau,\mu>$ of $\bar{P}_1$ in the previous example, action $<\tau,\mu>$ of $\bar{P}_3$ is followed by a choice between two exponentially timed τ-actions. It turns out that $\bar{P}_3 \approx_{\rm MB} \bar{P}_4$ because:

$$\begin{array}{rcl} pbtm(\bar{P}_3,[Q_1]_{\approx_{\rm MB}}) &=& \{\!|\, \frac{\gamma_1}{\gamma_1+\gamma_2}\cdot\left(\frac{1}{\mu}+\frac{1}{\gamma_1+\gamma_2}\right) \,|\!\} \;=\; pbtm(\bar{P}_4,[Q_1]_{\approx_{\rm MB}}) \\ pbtm(\bar{P}_3,[Q_2]_{\approx_{\rm MB}}) &=& \{\!|\, \frac{\gamma_2}{\gamma_1+\gamma_2}\cdot\left(\frac{1}{\mu}+\frac{1}{\gamma_1+\gamma_2}\right) \,|\!\} \;=\; pbtm(\bar{P}_4,[Q_2]_{\approx_{\rm MB}}) \end{array}$$

In general, for $n \in \mathbb{N}_{\ge 1}$ we have that $<\tau,\mu>.(<\tau,\gamma_1>.Q_1 + \dots + <\tau,\gamma_n>.Q_n)$ is weakly Markovian bisimilar to $<\tau,\frac{\gamma_1}{\gamma_1+\dots+\gamma_n}\cdot\left(\frac{1}{\mu}+\frac{1}{\gamma_1+\dots+\gamma_n}\right)^{-1}>.Q_1 + \dots + <\tau,\frac{\gamma_n}{\gamma_1+\dots+\gamma_n}\cdot\left(\frac{1}{\mu}+\frac{1}{\gamma_1+\dots+\gamma_n}\right)^{-1}>.Q_n$. ■

Example 3.5. Consider the following two process terms:

$$\begin{array}{l} \bar{P}_5 \;\equiv\; <\tau,\mu_1>.<\tau,\gamma>.Q_1 + <\tau,\mu_2>.<\tau,\gamma>.Q_2 \\ \bar{P}_6 \;\equiv\; <\tau,\frac{\mu_1}{\mu_1+\mu_2}\cdot\left(\frac{1}{\mu_1+\mu_2}+\frac{1}{\gamma}\right)^{-1}>.Q_1 + <\tau,\frac{\mu_2}{\mu_1+\mu_2}\cdot\left(\frac{1}{\mu_1+\mu_2}+\frac{1}{\gamma}\right)^{-1}>.Q_2 \end{array}$$

with $Q_1, Q_2 \in \mathbb{P}_{\rm nfu}$ and $Q_1 \not\approx_{\rm MB} Q_2$ as before. Unlike $\bar{P}_1$ and $\bar{P}_3$ in the previous two examples, $\bar{P}_5$ starts with a choice between two exponentially timed τ-actions, each of which is followed by the same action $<\tau,\gamma>$. It turns out that $\bar{P}_5 \approx_{\rm MB} \bar{P}_6$ because:

$$\begin{array}{rcl} pbtm(\bar{P}_5,[Q_1]_{\approx_{\rm MB}}) &=& \{\!|\, \frac{\mu_1}{\mu_1+\mu_2}\cdot\left(\frac{1}{\mu_1+\mu_2}+\frac{1}{\gamma}\right) \,|\!\} \;=\; pbtm(\bar{P}_6,[Q_1]_{\approx_{\rm MB}}) \\ pbtm(\bar{P}_5,[Q_2]_{\approx_{\rm MB}}) &=& \{\!|\, \frac{\mu_2}{\mu_1+\mu_2}\cdot\left(\frac{1}{\mu_1+\mu_2}+\frac{1}{\gamma}\right) \,|\!\} \;=\; pbtm(\bar{P}_6,[Q_2]_{\approx_{\rm MB}}) \end{array}$$

In general, for $n \in \mathbb{N}_{\ge 1}$ we have that $<\tau,\mu_1>.<\tau,\gamma>.Q_1 + \dots + <\tau,\mu_n>.<\tau,\gamma>.Q_n$ is weakly Markovian bisimilar to $<\tau,\frac{\mu_1}{\mu_1+\dots+\mu_n}\cdot\left(\frac{1}{\mu_1+\dots+\mu_n}+\frac{1}{\gamma}\right)^{-1}>.Q_1 + \dots + <\tau,\frac{\mu_n}{\mu_1+\dots+\mu_n}\cdot\left(\frac{1}{\mu_1+\dots+\mu_n}+\frac{1}{\gamma}\right)^{-1}>.Q_n$. The equivalence holds even if the derivative terms of actions $<\tau,\mu_i>$, $1 \le i \le n$, start with a choice among several exponentially timed τ-actions instead of a single exponentially timed τ-action, provided that all these derivative terms have the same total exit rate γ. ■

Example 3.6. We now examine all possible variants of $\bar{P}_5$ related to actions $<\tau, \gamma>$ and we show that none of these variants allows for any reduction, because it is not possible to preserve execution probabilities or expected durations of reducible computations. Firstly, consider the following two process terms:

$$\bar{P}_7 \equiv <\tau, \mu_1>.<\tau, \gamma_1>.Q_1 + <\tau, \mu_2>.<\tau, \gamma_2>.Q_2$$
$$\bar{P}_8 \equiv <\tau, \frac{\mu_1}{\mu_1+\mu_2} \cdot \left(\frac{1}{\mu_1+\mu_2} + \frac{1}{\gamma_1}\right)^{-1}>.Q_1 + <\tau, \frac{\mu_2}{\mu_1+\mu_2} \cdot \left(\frac{1}{\mu_1+\mu_2} + \frac{1}{\gamma_2}\right)^{-1}>.Q_2$$

where the two rates γ_1 and γ_2 are different from each other. Then $\bar{P}_7 \not\approx_{\rm MB} \bar{P}_8$ because for instance:

$$pbtm(\bar{P}_7, [Q_1]_{\approx_{\rm MB}}) = \{\!| \frac{\mu_1}{\mu_1+\mu_2} \cdot \left(\frac{1}{\mu_1+\mu_2} + \frac{1}{\gamma_1}\right) |\!\}$$
$$pbtm(\bar{P}_8, [Q_1]_{\approx_{\rm MB}}) = \{\!| \frac{\frac{\mu_1}{\mu_1+\mu_2} \cdot \left(\frac{1}{\mu_1+\mu_2} + \frac{1}{\gamma_1}\right)^{-1}}{\frac{\mu_1}{\mu_1+\mu_2} \cdot \left(\frac{1}{\mu_1+\mu_2} + \frac{1}{\gamma_1}\right)^{-1} + \frac{\mu_2}{\mu_1+\mu_2} \cdot \left(\frac{1}{\mu_1+\mu_2} + \frac{1}{\gamma_2}\right)^{-1}} \cdot \frac{1}{\frac{\mu_1}{\mu_1+\mu_2} \cdot \left(\frac{1}{\mu_1+\mu_2} + \frac{1}{\gamma_1}\right)^{-1} + \frac{\mu_2}{\mu_1+\mu_2} \cdot \left(\frac{1}{\mu_1+\mu_2} + \frac{1}{\gamma_2}\right)^{-1}} |\!\}$$

Secondly, consider the following two process terms:

$$\bar{P}_9 \equiv <\tau, \mu_1>.<\tau, \gamma>.Q_1 + <\tau, \mu_2>.Q_2$$
$$\bar{P}_{10} \equiv <\tau, \frac{\mu_1}{\mu_1+\mu_2} \cdot \left(\frac{1}{\mu_1+\mu_2} + \frac{1}{\gamma}\right)^{-1}>.Q_1 + <\tau, \mu_2>.Q_2$$

where the reduction takes place only in one of the two branches (equivalently, we could have taken $\bar{P}_9$ identical to $\bar{P}_5$ and modified $\bar{P}_{10}$ by inserting $<\tau, \gamma>$ after $<\tau, \mu_2>$). Then $\bar{P}_9 \not\approx_{\rm MB} \bar{P}_{10}$ because for instance:

$$pbtm(\bar{P}_9, [Q_2]_{\approx_{\rm MB}}) = \{\!| \frac{\mu_2}{\mu_1+\mu_2} \cdot \frac{1}{\mu_1+\mu_2} |\!\}$$
$$pbtm(\bar{P}_{10}, [Q_2]_{\approx_{\rm MB}}) = \{\!| \frac{\mu_2}{\frac{\mu_1}{\mu_1+\mu_2} \cdot \left(\frac{1}{\mu_1+\mu_2} + \frac{1}{\gamma}\right)^{-1} + \mu_2} \cdot \frac{1}{\frac{\mu_1}{\mu_1+\mu_2} \cdot \left(\frac{1}{\mu_1+\mu_2} + \frac{1}{\gamma}\right)^{-1} + \mu_2} |\!\}$$

■

All the reductions shown in Exs. 3.3, 3.4, and 3.5 can be summarized as follows.

Proposition 3.7. Let $I \neq \emptyset$ be a finite set, $J_i \neq \emptyset$ be a finite set for all $i \in I$, and $P_{i,j} \in \mathbb{P}$ for all $i \in I$ and $j \in J_i$. Then:

$$\sum_{i \in I} <\tau, \mu_i>. \sum_{j \in J_i} <\tau, \gamma_{i,j}>.P_{i,j} \approx_{\rm MB} \sum_{i \in I} \sum_{j \in J_i} <\tau, \frac{\mu_i}{\sum\limits_{k \in I} \mu_k} \cdot \frac{\gamma_{i,j}}{\sum\limits_{h \in J_i} \gamma_{i,h}} \cdot \left(\frac{1}{\sum\limits_{k \in I} \mu_k} + \frac{1}{\sum\limits_{h \in J_i} \gamma_{i,h}}\right)^{-1}>.P_{i,j}$$

whenever $\sum_{j \in J_{i_1}} \gamma_{i_1,j} = \sum_{j \in J_{i_2}} \gamma_{i_2,j}$ for all $i_1, i_2 \in I$. ■

3.2. Congruence Property

A desirable property of a behavioral equivalence is that of being a congruence with respect to the typical behavioral operators of process algebraic languages, thus enabling compositional reasoning. This is achieved by $\approx_{\rm MB}$ in the case of action prefix and hiding.

Proposition 3.8. Let $P_1, P_2 \in \mathbb{P}$. If $P_1 \approx_{\rm MB} P_2$, then:

1. $<a, \lambda>.P_1 \approx_{\rm MB} <a, \lambda>.P_2$ for all $<a, \lambda> \in Act$ (when $P_1, P_2 \in \mathbb{P}_{\rm seq}$).
2. $P_1 / H \approx_{\rm MB} P_2 / H$ for all $H \subseteq Name_{\rm v}$. ■

Relation $\approx_{\rm MB}$ is not a congruence with respect to the alternative composition operator. The problem has to do with fully unstable process terms. For instance, it holds that:

$$<\tau, \mu>.<\tau, \gamma>.\underline{0} \approx_{\rm MB} <\tau, \frac{\mu \cdot \gamma}{\mu+\gamma}>.\underline{0}$$

but:

$$<\tau, \mu>.<\tau, \gamma>.\underline{0} + <a, \lambda>.\underline{0} \not\approx_{\rm MB} <\tau, \frac{\mu \cdot \gamma}{\mu+\gamma}>.\underline{0} + <a, \lambda>.\underline{0}$$

In fact, if it were $a \neq \tau$, then we would have:

$$rate(<\tau, \mu>.<\tau, \gamma>.\underline{0} + <a, \lambda>.\underline{0},\ \tau,\ [\underline{0}]_{\approx_{\rm MB}}) = 0$$
$$rate(<\tau, \frac{\mu \cdot \gamma}{\mu+\gamma}>.\underline{0} + <a, \lambda>.\underline{0},\ \tau,\ [\underline{0}]_{\approx_{\rm MB}}) = \frac{\mu \cdot \gamma}{\mu+\gamma}$$

otherwise for $a = \tau$ we would have:

$$pbtm(<\tau, \mu>.<\tau, \gamma>.\underline{0} + <a, \lambda>.\underline{0},\ [\underline{0}]_{\approx_{\rm MB}}) = \{\!| \frac{\mu}{\mu+\lambda} \cdot \left(\frac{1}{\mu+\lambda} + \frac{1}{\gamma}\right), \frac{\lambda}{\mu+\lambda} \cdot \frac{1}{\mu+\lambda} |\!\}$$
$$pbtm(<\tau, \frac{\mu \cdot \gamma}{\mu+\gamma}>.\underline{0} + <a, \lambda>.\underline{0},\ [\underline{0}]_{\approx_{\rm MB}}) = \{\!| \frac{1}{\frac{\mu \cdot \gamma}{\mu+\gamma} + \lambda} |\!\}$$

The congruence violation with respect to the alternative composition operator can be prevented by adopting a construction analogous to the one used in [27] for weak bisimilarity over nondeterministic process terms. In other words, we have to apply the exit rate equality check also to fully unstable process terms, with the equivalence classes to consider being the ones with respect to $\approx_{\mathrm{MB}}$.

Definition 3.9. Let $P_1, P_2 \in \mathbb{P}$. We say that P_1 is *weakly Markovian bisimulation congruent* to P_2, written $P_1 \simeq_{\mathrm{MB}} P_2$, iff for all action names $a \in \mathit{Name}$ and equivalence classes $D \in \mathbb{P}/\approx_{\mathrm{MB}}$:

$$\mathit{rate}(P_1, a, D) \; = \; \mathit{rate}(P_2, a, D)$$

■

Proposition 3.10. $\sim_{\mathrm{MB}} \subsetneq \simeq_{\mathrm{MB}} \subsetneq \approx_{\mathrm{MB}}$, with $\simeq_{\mathrm{MB}} = \approx_{\mathrm{MB}}$ over $\mathbb{P}_{\mathrm{nfu}}$. ■

The following result is a straightforward consequence of the definition of $\simeq_{\mathrm{MB}}$.

Proposition 3.11. Let $P_1, P_2 \in \mathbb{P}_{\mathrm{seq}}$ and $<a, \lambda> \in \mathit{Act}$. Then $<a, \lambda>.P_1 \simeq_{\mathrm{MB}} <a, \lambda>.P_2$ iff $P_1 \approx_{\mathrm{MB}} P_2$. ■

It turns out that $\simeq_{\mathrm{MB}}$ is a congruence with respect to all the operators of MPC except for parallel composition (a counterexample will be shown at the beginning of Sect. 4).

Theorem 3.12. Let $P_1, P_2 \in \mathbb{P}$. If $P_1 \simeq_{\mathrm{MB}} P_2$, then:

1. $<a, \lambda>.P_1 \simeq_{\mathrm{MB}} <a, \lambda>.P_2$ for all $<a, \lambda> \in \mathit{Act}$ (when $P_1, P_2 \in \mathbb{P}_{\mathrm{seq}}$).
2. $P_1 + P \simeq_{\mathrm{MB}} P_2 + P$ and $P + P_1 \simeq_{\mathrm{MB}} P + P_2$ for all $P \in \mathbb{P}_{\mathrm{seq}}$ (when $P_1, P_2 \in \mathbb{P}_{\mathrm{seq}}$).
3. $P_1/H \simeq_{\mathrm{MB}} P_2/H$ for all $H \subseteq \mathit{Name}_{\mathrm{v}}$. ■

Moreover, $\simeq_{\mathrm{MB}}$ is the coarsest congruence, with respect to alternative composition, contained in $\approx_{\mathrm{MB}}$.

Theorem 3.13. Let $P_1, P_2 \in \mathbb{P}_{\mathrm{seq}}$. Then $P_1 \simeq_{\mathrm{MB}} P_2$ iff $P_1 + P \approx_{\mathrm{MB}} P_2 + P$ for all $P \in \mathbb{P}_{\mathrm{seq}}$. ■

Finally, $\simeq_{\mathrm{MB}}$ is a congruence also with respect to recursion. To show this, we need to extend $\simeq_{\mathrm{MB}}$ to open process terms in the usual way. The congruence proof is based on a notion of weak Markovian bisimulation up to $\approx_{\mathrm{MB}}$ inspired by the notion of Markovian bisimulation up to $\sim_{\mathrm{MB}}$ of [15]. It differs from its nondeterministic counterpart [27] due to the necessity, in this Markovian setting, of working with equivalence classes. In the following, we denote by $^+$ the operation of transitive closure for relations.

Definition 3.14. Let $P_1, P_2 \in \mathcal{PL}$ be process terms containing free occurrences of $k \in \mathbb{N}$ process variables $X_1, \ldots, X_k \in \mathit{Var}$ at most. We define $P_1 \simeq_{\mathrm{MB}} P_2$ iff $P_1\{Q_i \hookrightarrow X_i \mid 1 \le i \le k\} \simeq_{\mathrm{MB}} P_2\{Q_i \hookrightarrow X_i \mid 1 \le i \le k\}$ for all $Q_1, \ldots, Q_k \in \mathcal{PL}$ containing no free occurrences of process variables. ■

Definition 3.15. A binary relation $\mathcal{B}$ over $\mathbb{P}$ is a *weak Markovian bisimulation up to* $\approx_{\mathrm{MB}}$ iff, whenever $(P_1, P_2) \in \mathcal{B}$, then one of the following holds:

- $P_1, P_2 \in \mathbb{P}_{\mathrm{nfu}}$ and for all $a \in \mathit{Name}$ and $D \in \mathbb{P}/(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB}})^+$:
$$\mathit{rate}(P_1, a, D) \; = \; \mathit{rate}(P_2, a, D)$$
- $P_1, P_2 \in \mathbb{P}_{\mathrm{fu}}$ and for all $D \in \mathbb{P}_{\mathrm{nfu}}/(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB}})^+$:
$$\mathit{pbtm}(P_1, D) \; = \; \mathit{pbtm}(P_2, D)$$

■

Proposition 3.16. Let $\mathcal{B}$ be a binary relation over $\mathbb{P}$. If $\mathcal{B}$ is a weak Markovian bisimulation up to $\approx_{\mathrm{MB}}$, then $(P_1, P_2) \in \mathcal{B}$ implies $P_1 \approx_{\mathrm{MB}} P_2$ for all $P_1, P_2 \in \mathbb{P}$. Moreover $(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB}})^+ = \approx_{\mathrm{MB}}$. ■

Theorem 3.17. Let $P_1, P_2 \in \mathcal{PL}$ be sequential process terms containing free occurrences of $k \in \mathbb{N}$ process variables $X_1, \ldots, X_k \in \mathit{Var}$ at most. Whenever $P_1 \simeq_{\mathrm{MB}} P_2$, then

$$\mathrm{rec}\, X_1 : \cdots : \mathrm{rec}\, X_k : P_1 \;\simeq_{\mathrm{MB}}\; \mathrm{rec}\, X_1 : \cdots : \mathrm{rec}\, X_k : P_2$$

■

$$
\begin{array}{llcl}
(\mathcal{A}_{\mathrm{MB},1}) & P_1 + P_2 & = & P_2 + P_1 \\
(\mathcal{A}_{\mathrm{MB},2}) & (P_1 + P_2) + P_3 & = & P_1 + (P_2 + P_3) \\
(\mathcal{A}_{\mathrm{MB},3}) & P + \underline{0} & = & P \\
(\mathcal{A}_{\mathrm{MB},4}) & <a, \lambda_1>.P + <a, \lambda_2>.P & = & <a, \lambda_1 + \lambda_2>.P \\
(\mathcal{A}_{\mathrm{MB},5}) & <a, \lambda>. \sum\limits_{i \in I} <\tau, \mu_i>. \sum\limits_{j \in J_i} <\tau, \gamma_{i,j}>.P_{i,j} & = & <a, \lambda>. \sum\limits_{i \in I} \sum\limits_{j \in J_i} <\tau, \frac{\mu_i}{\mu} \cdot \frac{\gamma_{i,j}}{\gamma} \cdot \left(\frac{1}{\mu} + \frac{1}{\gamma} \right)^{-1}>.P_{i,j} \\
 & & & \text{if: } I \neq \emptyset \text{ finite, } J_i \neq \emptyset \text{ finite for all } i \in I \\
 & & & \mu = \sum\limits_{i \in I} \mu_i, \ \gamma = \sum\limits_{j \in J_i} \gamma_{i,j} \text{ for all } i \in I
\end{array}
$$

Table 2: Sound and complete axioms for $\simeq_{\mathrm{MB}}$ over $\mathbb{P}_{\mathrm{seq,nr}}$

3.3. Sound and Complete Axioms for Dynamic Operators

The relation $\simeq_{\mathrm{MB}}$ has a sound and complete axiomatization over the set $\mathbb{P}_{\mathrm{seq,nr}}$ of nonrecursive sequential process terms of $\mathbb{P}$. The axiomatization elucidates the equational laws associated with the operational definition of the equivalence and, thanks to its soundness and completeness, provides an algebraic alternative characterization of the equivalence itself. As a consequence, the axioms can be used to syntactically manipulate process terms as if they were rewriting rules consistent with the equivalence. In the specific case of $\simeq_{\mathrm{MB}}$, this will be exploited in Sect. 3.4 to investigate the exactness of the induced pseudo-aggregation.

As done at the end of Sect. 2 for $\sim_{\mathrm{MB}}$, also here we concentrate on the axioms for dynamic operators. The reason is that these axioms constitute the core of the equational characterization of an operationally defined behavioral equivalence. Moreover, the format of the axioms for other typical operators with respect to which the equivalence is a congruence – such as distribution laws for hiding or, alternatively, unfolding laws for recursion – is standard up to minor variations.

The axioms for $\simeq_{\mathrm{MB}}$ over $\mathbb{P}_{\mathrm{seq,nr}}$ are shown in Table 2. The first four axioms are inherited from $\sim_{\mathrm{MB}}$; they are valid for $\simeq_{\mathrm{MB}}$ too because $\sim_{\mathrm{MB}} \subsetneq \simeq_{\mathrm{MB}}$ as stated by Prop. 3.10. The fifth axiom schema characterizes $\simeq_{\mathrm{MB}}$; its validity comes from Props. 3.7 and 3.11. Standard inference rules for $\simeq_{\mathrm{MB}}$ expressing reflexivity, symmetry, transitivity, and substitutivity with respect to dynamic operators are kept implicit.

To prove completeness, we show that every process term in $\mathbb{P}_{\mathrm{seq,nr}}$ can be transformed into a normal form that abstracts from the order of summands (consistent with the first two axioms), rules out all null summands (consistent with the third axiom), and does not allow for simplifications based on the fourth and fifth axioms. Unlike the case of nondeterministic process terms, where saturation[1] is encoded in the normal form [27], here we cannot proceed that way, otherwise we would alter the quantitative behavior of the considered terms. In contrast, we elaborate on the result of Prop. 3.7 so to discover that pairs of terms related by $\approx_{\mathrm{MB}}$, but not by $\simeq_{\mathrm{MB}}$, enjoy one of two properties concerned with $\mathcal{A}_{\mathrm{MB},4}$ and $\mathcal{A}_{\mathrm{MB},5}$, respectively.

Lemma 3.18. Let $P_1, P_2 \in \mathbb{P}_{\mathrm{seq,nr}}$. If $P_1 \approx_{\mathrm{MB}} P_2$, but $P_1 \not\simeq_{\mathrm{MB}} P_2$, then P_1 and P_2 are respectively of the form:

$$\sum_{i \in I_1} <\tau, \mu_{1,i}>.P_{1,i} \quad \text{and} \quad \sum_{i \in I_2} <\tau, \mu_{2,i}>.P_{2,i}$$

where $I_1 \neq \emptyset, I_2 \neq \emptyset$ are finite index sets and at least one process term in $\{P_{1,i} \mid i \in I_1\} \cup \{P_{2,i} \mid i \in I_2\}$ is fully unstable. Moreover:

$$\{D \in \mathbb{P}/\!\approx_{\mathrm{MB}} \mid \exists i \in I_1.\, P_{1,i} \in D\} \ \neq \ \{D \in \mathbb{P}/\!\approx_{\mathrm{MB}} \mid \exists i \in I_2.\, P_{2,i} \in D\} \qquad \blacksquare$$

[1]A process term P is saturated iff, whenever P can reach P' after performing an a-action preceded and/or followed by finitely many τ-actions, then P can also reach P' directly through an a-action with no intervening τ-actions.

Proposition 3.19. Let $P_1, P_2 \in \mathbb{P}_{\rm seq,nr}$. If $P_1 \approx_{\rm MB} P_2$, but $P_1 \not\simeq_{\rm MB} P_2$, then at least one of P_1 and P_2 is of the form:

$$\sum_{i \in I} {<}\tau, \mu_i{>}. \sum_{j \in J_i} {<}\tau, \gamma_{i,j}{>}.P_{i,j}$$

where $I \neq \emptyset$ is a finite index set, $J_i \neq \emptyset$ is a finite index set for all $i \in I$, and one of the following two properties holds:

- $\sum\limits_{j \in J_{i_1}} {<}\tau, \gamma_{i_1,j}{>}.P_{i_1,j} \approx_{\rm MB} \sum\limits_{j \in J_{i_2}} {<}\tau, \gamma_{i_2,j}{>}.P_{i_2,j}$ for all $i_1, i_2 \in I$.
- $\sum\limits_{j \in J_{i_1}} \gamma_{i_1,j} = \sum\limits_{j \in J_{i_2}} \gamma_{i_2,j}$ for all $i_1, i_2 \in I$. ■

Definition 3.20. We say that $P \in \mathbb{P}_{\rm seq,nr}$ is in $\simeq_{\rm MB}$-*normal-form* iff either P is $\underline{0}$, or P is of the form $\sum_{i \in I} {<}a_i, \lambda_i{>}.P_i$ with I finite and nonempty, P initially minimal with respect to $\mathcal{A}_{\rm MB,4}$, ${<}a_i, \lambda_i{>}.P_i$ initially minimal with respect to $\mathcal{A}_{\rm MB,5}$ for all $i \in I$, and P_i in $\simeq_{\rm MB}$-normal-form for all $i \in I$. ■

In the definition above, by P initially minimal with respect to $\mathcal{A}_{\rm MB,4}$ we mean that P does not contain any two summands like the ones on the left-hand side of $\mathcal{A}_{\rm MB,4}$. Likewise, by ${<}a_i, \lambda_i{>}.P_i$ initially minimal with respect to $\mathcal{A}_{\rm MB,5}$ we mean that ${<}a_i, \lambda_i{>}.P_i$ does not match the left-hand side of $\mathcal{A}_{\rm MB,5}$.

It is worth noting that, by virtue of Prop. 3.19, whenever it holds that $P_1 \approx_{\rm MB} P_2$, but $P_1 \not\simeq_{\rm MB} P_2$, then at least one of P_1 and P_2 is not in $\simeq_{\rm MB}$-normal-form because of a violation of initial minimality with respect to $\mathcal{A}_{\rm MB,4}$ or $\mathcal{A}_{\rm MB,5}$. This will be exploited in the proof of the completeness part of Thm. 3.22 below.

Lemma 3.21. For each $P \in \mathbb{P}_{\rm seq,nr}$ there exists $Q \in \mathbb{P}_{\rm seq,nr}$ in $\simeq_{\rm MB}$-normal-form such that $\mathcal{A}_{\rm MB} \vdash P = Q$. ■

Theorem 3.22. Let $P_1, P_2 \in \mathbb{P}_{\rm seq,nr}$. Then $\mathcal{A}_{\rm MB} \vdash P_1 = P_2 \iff P_1 \simeq_{\rm MB} P_2$. ■

3.4. Exactness at Stationary State

Weak Markovian bisimilarity $\approx_{\rm MB}$ and the coarsest congruence contained in it, $\simeq_{\rm MB}$, are more liberal than Markovian bisimilarity $\sim_{\rm MB}$, because they allow every sequence of exponentially timed τ-transitions to be considered equivalent to a single exponentially timed τ-transition having the same expected duration. From a stochastic viewpoint, this amounts to approximating a hypoexponentially (or Erlang) distributed random variable with an exponentially distributed random variable having the same expected value. From a performance evaluation viewpoint, this can be exploited to assess more quickly properties expressed in terms of the mean time to certain events by working on an aggregated CTMC. To be precise, since the Markov property of the original CTMC is not preserved after the approximation, but the stochastic process resulting from the aggregation is still assumed to be a CTMC, it is more appropriate to call the aggregation a *pseudo-aggregation* [32].

Fortunately, mean-time-to properties are not the only ones preserved by the two weak Markovian behavioral equivalences that we have introduced. Indeed, we will see that the CTMC-level pseudo-aggregation induced by such equivalences is *exact at stationary state*. As a consequence, the two weak Markovian behavioral equivalences can be used for reducing the size of models possessing an underlying CTMC-based semantics without altering the value of reward-based [23] performance measures at stationary state. In general, this is true as long as states (resp. transitions) to be aggregated are given equal cumulative rewards (resp. appropriate instantaneous rewards), otherwise reward-sensitive refinements of $\approx_{\rm MB}$ and $\simeq_{\rm MB}$ inspired by [13] should be employed. In this weak setting, it is sufficient that rewards are associated neither with fully unstable states, nor with exponentially timed τ-transitions, which is quite reasonable.

Usually, CTMC-level aggregations are defined in a *structural* way, i.e., in terms of the conditions that CTMC states and transitions have to fulfill. This was the case for ordinary lumpability, which is expressed in the same way as Markovian bisimilarity [22, 16] up to the fact that there are no action labels on CTMC transitions (which is equivalent to viewing all CTMC transitions as being labeled with the same action). A different approach was taken to formalize T-lumpability, the aggregation induced by Markovian testing and trace equivalences [8]. The idea was that of exploiting the sound and complete axiomatization of

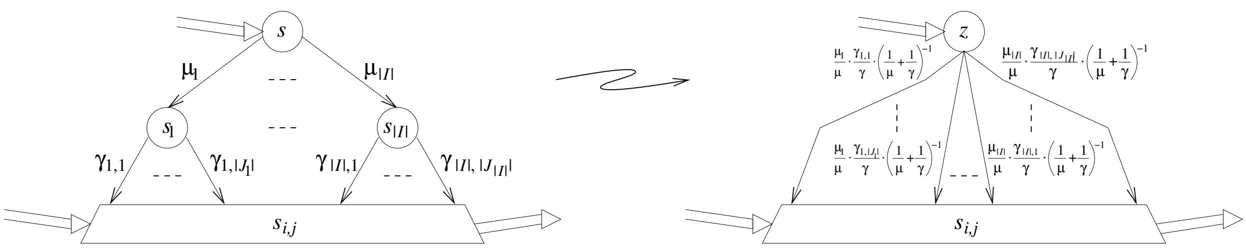

Figure 2: Effect of W-lumpability derived from axiom $\mathcal{A}_{\mathrm{MB},5}$ ($\mu = \sum_{i \in I} \mu_i$, $\gamma = \sum_{j \in J_i} \gamma_{i,j}$ for all $i \in I$)

Markovian testing equivalence, in particular the axiom differentiating this equivalence from Markovian bisimilarity. Subsequently, a structural characterization of T-lumpability was given in [33] through a notion called weighted lumpability that takes a two-step perspective. Unlike ordinary lumpability, states are not compared on the basis of the rates with which they reach their direct successors. Rather, the weighted rates with which they reach their successors at distance two are considered.

The pseudo-aggregation induced by $\approx_{\mathrm{MB}}$ and $\simeq_{\mathrm{MB}}$ can be investigated by exploiting the sound and complete axiomatization of Sect. 3.3, in particular the characterizing axiom $\mathcal{A}_{\mathrm{MB},5}$ without its two initial $<a, \lambda>$ actions (which are not necessary in the case of $\approx_{\mathrm{MB}}$). If we view this axiom as the rewriting rule shown in Fig. 2, then we can think of a CTMC as being W-lumpable iff a portion of its state space matches the left-hand side of the rewriting rule, in which case it is replaced by the right-hand side where the topmost $1 + |I|$ states have been merged into a single one. Notice that states $s_{i,j}$ can have arbitrarily many incoming and outgoing transitions collectively depicted through a double arrow; similarly for the incoming transitions of s and z. In contrast, all the transitions departing from s and all the incoming and outgoing transitions of each of the states s_1 to $s_{|I|}$ are depicted in the figure. If some of the states s_1 to $s_{|I|}$ were reachable in one step also from a state different from s, it should be duplicated before the aggregation takes place.

Axiom $\mathcal{A}_{\mathrm{MB},5}$ and the related rewriting rule in Fig. 2 are instrumental to develop a two-step structural characterization of W-lumpability. In the following, we formalize a CTMC as a pair (S, R) where S is the set of states and $R : S \times S \to \mathbb{R}_{\geq 0}$ is the rate function, with $R(s, s') = 0$ meaning that s' is not reachable in one step from s. With abuse of notation, we let $R(s, S') = \sum_{s' \in S'} R(s, s')$ for $S' \subseteq S$. Moreover, we define the total exit rate of s as $E(s) = R(s, S)$.

Definition 3.23. Let (S, R) be a CTMC. An equivalence relation $\mathcal{R}$ over S is a *W-lumping* iff, whenever $(s_1, s_2) \in \mathcal{R}$, then one of the following three conditions holds:

- For all equivalence classes $D \in S/\mathcal{R}$:
$$R(s_1, D) \; = \; R(s_2, D)$$
- For one of s_1 and s_2, which we denote by s, there exist $\mu, \gamma \in \mathbb{R}_{>0}$ such that:
  - $E(s) = \mu$,
  - $E(s') = \gamma$ for all $s' \in S$ such that $R(s, s') > 0$,
  - $\{s'' \in S \mid R(s'', s') > 0\} = \{s\}$ for all $s' \in S$ such that $R(s, s') > 0$,

  while for the other one of s_1 and s_2, which we denote by z, it holds that:
  - $E(z) = \left(\frac{1}{\mu} + \frac{1}{\gamma}\right)^{-1}$.

  Moreover, for all equivalence classes $D \in S/\mathcal{R}$:
$$\frac{R(z,D)}{E(z)} \; = \sum_{D' \in S/\mathcal{R}} \left(\frac{R(s,D')}{E(s)} \cdot \sum_{s' \in D' \text{ s.t. } R(s,s')>0} \frac{R(s',D)}{E(s')} \right)$$

- For each $i = 1, 2$ there exist $\mu_i, \gamma_i \in \mathbb{R}_{>0}$ such that:
  - $E(s_i) = \mu_i$.
  - $E(s') = \gamma_i$ for all $s' \in S$ such that $R(s_i, s') > 0$.
  - $\{s'' \in S \mid R(s'', s') > 0\} = \{s_i\}$ for all $s' \in S$ such that $R(s_i, s') > 0$.

  Moreover:
$$\left(\frac{1}{\mu_1} + \frac{1}{\gamma_1}\right)^{-1} = \left(\frac{1}{\mu_2} + \frac{1}{\gamma_2}\right)^{-1}$$
  and for all equivalence classes $D \in S/\mathcal{R}$:
$$\sum_{D' \in S/\mathcal{R}} \left(\frac{R(s_1,D')}{E(s_1)} \cdot \sum_{s' \in D' \text{ s.t. } R(s_1,s')>0} \frac{R(s',D)}{E(s')} \right) = \sum_{D' \in S/\mathcal{R}} \left(\frac{R(s_2,D')}{E(s_2)} \cdot \sum_{s' \in D' \text{ s.t. } R(s_2,s')>0} \frac{R(s',D)}{E(s')} \right)$$

We say that $s_1, s_2 \in S$ are *W-lumpable* iff $(s_1, s_2) \in \mathcal{R}$ for some W-lumping $\mathcal{R}$. ■

Notice that the first condition above is ordinary lumpability, the second one corresponds to the rewriting rule in Fig. 2 (two-step reduction only on one side), and the third one manages situations in which a two-step reduction is needed on both sides (e.g., to identify $<\tau, \mu>.<\tau, \gamma>.Q$ and $<\tau, \gamma>.<\tau, \mu>.Q$). The following result shows that W-lumpability is exact at stationary state for all the processes in $\mathbb{P}$.

Theorem 3.24. W-lumpability is exact at stationary state, i.e., the stationary probability of being in a macrostate of a CTMC obtained via W-lumpability is the sum of the stationary probabilities of being in any of the constituent microstates of the CTMC from which the aggregated one has been obtained. ■

Stationary state exactness can be proved directly on the CTMC models at hand by transforming the global balance equations of the original CTMC into a form equivalent to the global balance equations of the aggregated CTMC. Unlike ordinary lumpability and T-lumpability, W-lumpability is not exact at transient state, hence properties expressed in terms of transient state probabilities may not be preserved.

Example 3.25. Consider again process terms $\bar{P}_1$ and $\bar{P}_2$ of Ex. 3.3 (see also the leftmost part of Fig. 1). The sum of the probabilities of being in one of the first two states of $[\![\bar{P}_1]\!]$ at time $t \in \mathbb{R}_{>0}$ is different from the probability of being in the first state of $[\![\bar{P}_2]\!]$ at the same time instant. In fact, the probability of being in that state of $[\![\bar{P}_2]\!]$ at that time is the probability that the exponentially distributed duration of its outgoing transition is greater than t, which is $1 - (1 - \mathrm{e}^{-\frac{\mu \cdot \gamma}{\mu + \gamma} \cdot t}) = \mathrm{e}^{-\frac{\mu \cdot \gamma}{\mu + \gamma} \cdot t}$ and reduces to $\mathrm{e}^{-\frac{\mu}{2} \cdot t}$ when $\mu = \gamma$. In contrast, the probability of being in one of those states of $[\![\bar{P}_1]\!]$ at that time is the probability that the hypoexponentially (for $\mu \neq \gamma$) or Erlang (for $\mu = \gamma$) distributed duration of their two consecutive outgoing transitions is greater than t, which is $1 - (1 - \frac{\gamma}{\gamma - \mu} \cdot \mathrm{e}^{-\mu \cdot t} + \frac{\mu}{\gamma - \mu} \cdot \mathrm{e}^{-\gamma \cdot t}) = \frac{\gamma}{\gamma - \mu} \cdot \mathrm{e}^{-\mu \cdot t} - \frac{\mu}{\gamma - \mu} \cdot \mathrm{e}^{-\gamma \cdot t}$ or $1 - (1 - (1 + \mu \cdot t) \cdot \mathrm{e}^{-\mu \cdot t}) = (1 + \mu \cdot t) \cdot \mathrm{e}^{-\mu \cdot t}$, respectively. ■

3.5. Coarser Weak Markovian Bisimilarities

In Sect. 3.1, we have defined $pbtm(P, D)$ as the t-indexed multiset of sums of *probtime* measures over all reducible computations from P to D *having the same expected duration* t. If there are $n \in \mathbb{N}_{\geq 1}$ reducible computations from P to D with expected duration t whose execution probabilities are $p_1, p_2, \ldots, p_n$, then the corresponding value in $pbtm(P, D)$ is simply $t \cdot \sum_{1 \leq i \leq n} p_i$.

As we have shown, taking these sums is necessary in order for $\approx_{\mathrm{MB}}$ to be a conservative extension of $\sim_{\mathrm{MB}}$. However, this is in some sense the *minimal* requirement for guaranteeing compatibility with $\sim_{\mathrm{MB}}$ over fully unstable process terms. A more radical option would be to add up the *probtime* measures of all the reducible computations from P to D *without considering their expected durations*. This amounts to replacing $pbtm(P, D)$ with $pbtm'(P, D)$ where:
$$pbtm'(P, D) = \sum_{t \in \mathbb{R}_{>0} \text{ s.t. } rcomp(P,D,t) \neq \emptyset} \; \sum_{c \in rcomp(P,D,t)} probtime(c)$$
and performing a $pbtm'$-based equality check in the definition of $\approx_{\mathrm{MB}}$, thereby obtaining a coarser weak Markovian bisimulation equivalence that we denote by $\approx'_{\mathrm{MB}}$. If all the reducible computations from P lead to D, then $pbtm'(P, D)$ is the expected time to reach D from P thanks to the way *probtime* is defined.

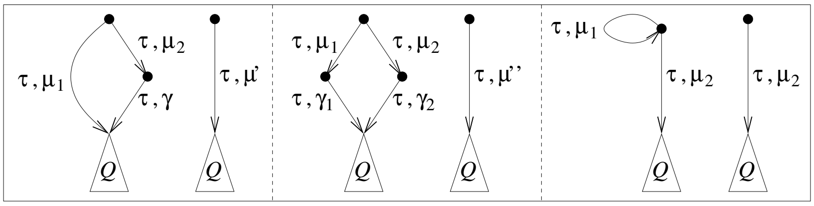

Figure 3: Additional identifications performed by $\approx'_{\rm MB}$ with respect to $\approx_{\rm MB}$ (see Ex. 3.26)

Example 3.26. Figure 3 shows further identifications made feasible by $\approx'_{\rm MB}$ thanks to the possibility of summing up the *probtime* measures of reducible computations (from the same source state to the same set of target states) having different expected durations:

- In the leftmost part of the figure – where the two reducible computations c_1 and c_2 to Q in the leftmost model have different expected durations – the identification holds iff:
$$\begin{aligned}\frac{1}{\mu'} &= probtime(c_1) + probtime(c_2) \\ &= \left(\frac{\mu_1}{\mu_1+\mu_2}\cdot\frac{1}{\mu_1+\mu_2}\right) + \left(\left(\frac{\mu_2}{\mu_1+\mu_2}\cdot\frac{\gamma}{\gamma}\right)\cdot\left(\frac{1}{\mu_1+\mu_2}+\frac{1}{\gamma}\right)\right) \\ &= \frac{1}{\mu_1+\mu_2} + \frac{\mu_2}{\gamma\cdot(\mu_1+\mu_2)}\end{aligned}$$
i.e.:
$$\mu' = \frac{\gamma\cdot(\mu_1+\mu_2)}{\gamma+\mu_2}$$
- In the middle part of the figure – where the two reducible computations c_1 and c_2 to Q in the leftmost model have different expected durations for $\gamma_1 \neq \gamma_2$, otherwise we are in a situation similar to Ex. 3.5 (see also the rightmost part of Fig. 1) – the identification holds iff:
$$\begin{aligned}\frac{1}{\mu''} &= probtime(c_1) + probtime(c_2) \\ &= \left(\left(\frac{\mu_1}{\mu_1+\mu_2}\cdot\frac{\gamma_1}{\gamma_1}\right)\cdot\left(\frac{1}{\mu_1+\mu_2}+\frac{1}{\gamma_1}\right)\right) + \left(\left(\frac{\mu_2}{\mu_1+\mu_2}\cdot\frac{\gamma_2}{\gamma_2}\right)\cdot\left(\frac{1}{\mu_1+\mu_2}+\frac{1}{\gamma_2}\right)\right) \\ &= \frac{1}{\mu_1+\mu_2} + \frac{\mu_1}{\gamma_1\cdot(\mu_1+\mu_2)} + \frac{\mu_2}{\gamma_2\cdot(\mu_1+\mu_2)}\end{aligned}$$
i.e.:
$$\mu'' = \frac{\gamma_1\cdot\gamma_2\cdot(\mu_1+\mu_2)}{\gamma_1\cdot\gamma_2+\mu_1\cdot\gamma_2+\mu_2\cdot\gamma_1}$$
- In the rightmost part of the figure – where there are countably many reducible computations c_n to Q in the leftmost model, each returning to the initial state $n \in \mathbb{N}$ times – the identification holds because:
$$\begin{aligned}\sum_{n=0}^{\infty} probtime(c_n) &= \sum_{n=0}^{\infty}\left(\left(\left(\frac{\mu_1}{\mu_1+\mu_2}\right)^n\cdot\frac{\mu_2}{\mu_1+\mu_2}\right)\cdot\left((n+1)\cdot\frac{1}{\mu_1+\mu_2}\right)\right) \\ &= \frac{\mu_2}{(\mu_1+\mu_2)^2}\cdot\left(\sum_{n=0}^{\infty} n\cdot\left(\frac{\mu_1}{\mu_1+\mu_2}\right)^n + \sum_{n=0}^{\infty}\left(\frac{\mu_1}{\mu_1+\mu_2}\right)^n\right) \\ &= \frac{\mu_2}{(\mu_1+\mu_2)^2}\cdot\left(\frac{\mu_1/(\mu_1+\mu_2)}{(1-\mu_1/(\mu_1+\mu_2))^2} + \frac{1}{(1-\mu_1/(\mu_1+\mu_2))}\right) \\ &= \frac{\mu_2}{(\mu_1+\mu_2)^2}\cdot\frac{1}{(1-\mu_1/(\mu_1+\mu_2))^2} \\ &= \frac{1}{\mu_2}\end{aligned}$$
This means that $\approx'_{\rm MB}$ can eliminate exponentially timed internal *selfloops*, which is consistent with the fact that selfloops are ignored when solving a CTMC. In general, *cycles* of exponentially timed internal transitions equipped with a way out (i.e., an exiting exponentially timed internal transition) can be eliminated by $\approx'_{\rm MB}$, but their rates cannot be neglected. In fact, if we slightly modify the leftmost model by transforming the exponentially timed internal selfloop insisting on the initial state into a transition labeled with $<\tau, \mu_1>$ to a new state followed by a transition labeled with $<\tau, \gamma>$ back to the initial state, then the rate μ''' of the initial transition of the rightmost model should satisfy:
$$\mu''' = \frac{\mu_2\cdot\gamma\cdot(\mu_1+\mu_2)}{\mu_2\cdot\gamma+\mu_1\cdot(\gamma+\mu_1+\mu_2)} \qquad \blacksquare$$

Relation $\approx'_{\rm MB}$ and the coarsest congruence $\simeq'_{\rm MB}$ contained in it enjoy properties analogous to those of relations $\approx_{\rm MB}$ and $\simeq'_{\rm MB}$, respectively. In particular, as far as the axiomatization over $\mathbb{P}_{\rm seq,nr}$ is concerned, what we have to add is an axiom schema like $\mathcal{A}_{\rm MB,5}$ where the various branches, which can now have different lengths, all lead to the same process term like in $\mathcal{A}_{\rm MB,4}$ (see the leftmost and middle parts of Fig. 3). The additional axiom schema is:

$$< a, \lambda >. \sum_{i_1 \in I_1} <\tau, \mu_{i_1}>. \sum_{i_2 \in I_{i_1,2}} <\tau, \mu_{i_1,i_2}>. \dots . \sum_{i_n \in I_{i_1,i_2,\dots,n}} <\tau, \mu_{i_1,i_2,\dots,i_n}>.P \;=\; < a, \lambda >.<\tau, \mu>.P$$

if:

$$\frac{1}{\mu} \;=\; \sum_{i_1 \in I_1} \sum_{i_2 \in I_{i_1,2}} \dots \sum_{i_n \in I_{i_1,i_2,\dots,n}} \frac{\mu_{i_1}}{M_1} \cdot \frac{\mu_{i_1,i_2}}{M_{i_1,2}} \cdot \dots \cdot \frac{\mu_{i_1,i_2,\dots,i_n}}{M_{i_1,i_2,\dots,n}} \cdot \left(\frac{1}{M_1} + \frac{1}{M_{i_1,2}} + \dots + \frac{1}{M_{i_1,i_2,\dots,n}} \right)$$

where $n \in \mathbb{N}_{\geq 1}$, $I_1 \neq \emptyset$, $M_{i_1,i_2,\dots,j} = \sum_{i_j \in I_{i_1,i_2,\dots,j}} \mu_{i_1,i_2,\dots,i_j}$ for all $1 \leq j \leq n$, and $I_{i_1,i_2,\dots,j} = \emptyset$ for some $2 \leq j \leq n$ means that the related summation term is P and the related contribution to $1/\mu$ is null.

4. Enhancing the Abstraction Capability: Full Compositionality, Limited Exactness

The relation $\simeq_{\rm MB}$ is not a congruence with respect to the parallel composition operator. This fact, which is illustrated below, restricts the usefulness of $\simeq_{\rm MB}$ for compositional state space reduction purposes in the framework of integrated-time Markovian process calculi.

Example 4.1. Assume that parallel composition has lower priority than any other operator. It holds that:

$$< a, \lambda >.<\tau, \mu>.<\tau, \gamma>.\underline{0} \;\simeq_{\rm MB}\; < a, \lambda >.<\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0}$$

while in the context of $\_ \,\|_{\emptyset}\, < a', \lambda' >.\underline{0}$ it turns out that:

$$< a, \lambda >.<\tau, \mu>.<\tau, \gamma>.\underline{0} \,\|_{\emptyset}\, < a', \lambda' >.\underline{0} \;\not\simeq_{\rm MB}\; < a, \lambda >.<\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0} \,\|_{\emptyset}\, < a', \lambda' >.\underline{0}$$

To show this, first of all we note that:

$$<\tau, \mu>.<\tau, \gamma>.\underline{0} \,\|_{\emptyset}\, < a', \lambda' >.\underline{0} \;\not\approx_{\rm MB}\; <\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0} \,\|_{\emptyset}\, < a', \lambda' >.\underline{0}$$

In fact, for $a' \neq \tau$ the two process terms are not fully unstable and it holds that:

$$\begin{aligned} \mathit{rate}(<\tau, \mu>.<\tau, \gamma>.\underline{0} \,\|_{\emptyset}\, < a', \lambda' >.\underline{0},\, \tau,\, [<\tau, \gamma>.\underline{0} \,\|_{\emptyset}\, < a', \lambda' >.\underline{0}]_{\approx_{\rm MB}}) \;&=\; \mu \\ \mathit{rate}(<\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0} \,\|_{\emptyset}\, < a', \lambda' >.\underline{0},\, \tau,\, [<\tau, \gamma>.\underline{0} \,\|_{\emptyset}\, < a', \lambda' >.\underline{0}]_{\approx_{\rm MB}}) \;&=\; 0 \end{aligned}$$

On the other hand, for $a' = \tau$ the two process terms are fully unstable and it holds that:

$$\begin{aligned} \mathit{pbtm}(<\tau, \mu>.<\tau, \gamma>.\underline{0} \,\|_{\emptyset}\, < a', \lambda' >.\underline{0},\, [\underline{0} \,\|_{\emptyset}\, \underline{0}]_{\approx_{\rm MB}}) \;=\; \{\!|\; & (\tfrac{\mu}{\mu + \lambda'} \cdot \tfrac{\gamma}{\gamma + \lambda'}) \cdot (\tfrac{1}{\mu + \lambda'} + \tfrac{1}{\gamma + \lambda'} + \tfrac{1}{\lambda'}), \\ & (\tfrac{\mu}{\mu + \lambda'} \cdot \tfrac{\lambda'}{\gamma + \lambda'}) \cdot (\tfrac{1}{\mu + \lambda'} + \tfrac{1}{\gamma + \lambda'} + \tfrac{1}{\gamma}), \\ & (\tfrac{\lambda'}{\mu + \lambda'}) \cdot (\tfrac{1}{\mu + \lambda'} + \tfrac{1}{\mu} + \tfrac{1}{\gamma}) \;|\!\} \end{aligned}$$

$$\begin{aligned} \mathit{pbtm}(<\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0} \,\|_{\emptyset}\, < a', \lambda' >.\underline{0},\, [\underline{0} \,\|_{\emptyset}\, \underline{0}]_{\approx_{\rm MB}}) \;=\; \{\!|\; & (\tfrac{\frac{\mu \cdot \gamma}{\mu + \gamma}}{\frac{\mu \cdot \gamma}{\mu + \gamma} + \lambda'}) \cdot (\tfrac{1}{\frac{\mu \cdot \gamma}{\mu + \gamma} + \lambda'} + \tfrac{1}{\lambda'}), \\ & (\tfrac{\lambda'}{\frac{\mu \cdot \gamma}{\mu + \gamma} + \lambda'}) \cdot (\tfrac{1}{\frac{\mu \cdot \gamma}{\mu + \gamma} + \lambda'} + \tfrac{1}{\frac{\mu \cdot \gamma}{\mu + \gamma}}) \;|\!\} \end{aligned}$$

Thus:

$$[<\tau, \mu>.<\tau, \gamma>.\underline{0} \,\|_{\emptyset}\, < a', \lambda' >.\underline{0}]_{\approx_{\rm MB}} \cap [<\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0} \,\|_{\emptyset}\, < a', \lambda' >.\underline{0}]_{\approx_{\rm MB}} \;=\; \emptyset$$

and, as a consequence, we have that:

$$\begin{aligned} \mathit{rate}(< a, \lambda >.<\tau, \mu>.<\tau, \gamma>.\underline{0} \,\|_{\emptyset}\, < a', \lambda' >.\underline{0},\, a,\, [<\tau, \mu>.<\tau, \gamma>.\underline{0} \,\|_{\emptyset}\, < a', \lambda' >.\underline{0}]_{\approx_{\rm MB}}) \;&=\; \lambda \\ \mathit{rate}(< a, \lambda >.<\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0} \,\|_{\emptyset}\, < a', \lambda' >.\underline{0},\, a,\, [<\tau, \mu>.<\tau, \gamma>.\underline{0} \,\|_{\emptyset}\, < a', \lambda' >.\underline{0}]_{\approx_{\rm MB}}) \;&=\; 0 \end{aligned}$$

Also the two divergent process terms $\mathrm{rec}\, X : <\tau, \mu>.<\tau, \gamma_1>.X$ and $\mathrm{rec}\, X : <\tau, \mu>.<\tau, \gamma_2>.X$, $\gamma_1 \neq \gamma_2$, are related by $\simeq_{\rm MB}$, but this no longer holds when placing them in the context of $\_ \,\|_{\emptyset}\, < a', \lambda' >.\underline{0}$ with $a' \neq \tau$. ■

In order to overcome the drawback exemplified above, in Sect. 4.1 we revise the notion of reducible computation and, based on this revision, in Sect. 4.2 we define a generalized weak variant of $\sim_{\rm MB}$ over MPC that enhances the abstraction capability of $\approx_{\rm MB}$ in the presence of interleaved computations. We then prove in Sect. 4.3 that the generalized equivalence is a congruence also with respect to parallel composition, but in Sect. 4.4 we show that it induces a pseudo-aggregation that is exact at stationary state only for a subset of processes satisfying a certain constraint on synchronizations. Finally, in Sect. 4.5 we outline coarser variants.

Figure 4: Reduction of acyclic replicated computations in a concurrent setting

4.1. Revising the Notion of Reducible Computation

As we have seen, $\approx_{\rm MB}$ and $\simeq_{\rm MB}$ abstract from sequences of exponentially timed τ-transitions while preserving (at the computation level) their execution probability and expected duration as well as (at the system level) transient properties expressed in terms of the mean time to certain events and stationary-state reward-based performance measures. This kind of abstraction has been done in the simplest possible case: sequences of exponentially timed τ-actions labeling computations that traverse *fully unstable states*.

When dealing with concurrent processes, a revision of the notion of reducible computation seems unavoidable to achieve compositionality. In this setting, we need to address the case of sequences of exponentially timed τ-actions labeling computations that traverse *unstable states satisfying certain conditions*. The reason is that, if we view a system description as the parallel composition of several sequential components, any of those components may have local computations traversing *fully unstable local states*, but in the overall system those local states may be *part of global states that are not fully unstable*.

For instance, this is the case with the process $<\tau, \mu>.<\tau, \gamma>.\underline{0} \,\|_{\emptyset}\, <a, \lambda>.\underline{0}$, whose underlying labeled multitransition system is depicted on the left-hand side of Fig. 4. As can be noted, the fully unstable local states traversed by the only local computation of the sequential component $<\tau, \mu>.<\tau, \gamma>.\underline{0}$ may become part of unstable global states that are not fully unstable if $a \neq \tau$. Our objective is to revise the notion of reducible computation in such a way that the labeled multitransition system on the left-hand side of Fig. 4 can be regarded as being weakly Markovian bisimilar to the labeled multitransition system on the right-hand side. Notice that this implies that execution probabilities and expected durations can only be preserved *at the level of local computations*, hence transient properties expressed in terms of the mean time to certain events may no longer be preserved at the system level.

In a concurrent setting, a sequence of exponentially timed τ-actions may be *replicated due to interleaving*, in the sense that it may label several computations that share no transition. The revision of the notion of reducible computation is thus based on the idea that, for each computation that traverses fully unstable *local* states and is labeled with exponentially timed τ-actions, we have to recognize *all the replicas* of that computation and pinpoint their *initial and final states*. On the left-hand side of Fig. 4, there are two replicas with initial states $s_{1,1}$ and $s_{1,4}$ and final states $s_{1,3}$ and $s_{1,6}$. In general, a one-to-one correspondence can be established between the states traversed by any two replicas by following the direction of the transitions. On the left-hand side of Fig. 4, the pairs of corresponding states are composed of the two initial states $(s_{1,1}, s_{1,4})$, the two intermediate states $(s_{1,2}, s_{1,5})$, and the two final states $(s_{1,3}, s_{1,6})$. Thus, we can say that, *when moving vertically, the current stage of the replicas is preserved.*

In addition to the exponentially timed τ-transition belonging to the replica, any two states traversed by the same replica can only possess transitions that are pairwise identically labeled. Those transitions are emanated from (the local states of) sequential components that are in parallel with (the local state of) the sequential component originating the considered reducible computation. The set of those transitions not belonging to the replica can thus be viewed as the *context of the replica.* On the left-hand side of Fig. 4, the context of the top replica has a single transition labeled with $<a, \lambda>$, whereas the context of the bottom replica is empty. Thus, *when moving horizontally, the context of each replica is preserved*, i.e., the context does not change along a replica. On the other hand, *different replicas may have different contexts.*

With regard to the identification of the boundary of the replicas of a reducible computation, there are two possibilities. One is that the final states have no exponentially timed τ-transitions, as on the left-hand side of Fig. 4. The other is that, at a certain point, each replica has an exponentially timed τ-transition back to one of the preceding states of the replica itself, as shown on the left-hand side of Fig. 5. In this case, for each replica we view its return state as being its final state. Thus, on the left-hand side of Fig. 5, for

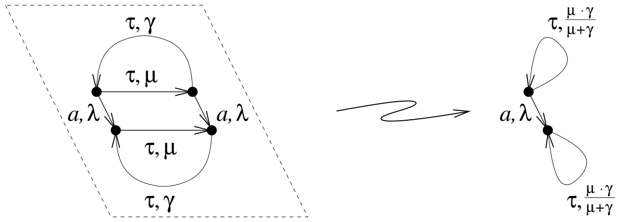

Figure 5: Reduction of cyclic replicated computations in a concurrent setting

both replicas the final state coincides with the initial state. In general, a cyclic reducible computation may result from the synchronization of cyclic computations of several sequential components, with the lengths of the various cycles being possibly different. In that case, we take as final state of the reducible computation the one that is reached after a number of steps equal to the length of the longest cycle.

All the considerations above lead us to proceed as follows. Firstly, consistent with the two-level syntax in Def. 2.1, we adopt the view that a process term represents a system made out of several components. The state corresponding to a possibly concurrent process term P can thus be viewed as a vector of ℓ local states, each corresponding to a sequential component C occurring in a different position $k = 1, \dots, \ell$ of P.

Secondly, we revise the notion of reducibility in such a way that it applies only to computations of (fully synchronized portions of) sequential components, whose length is maximal up to cycles. For a correct account of fully unstable and non-fully-unstable states, the behavior of a sequential component $C \in \mathbb{P}_{\mathrm{seq}}$ occurring in position k of $P \in \mathbb{P}$ should be *considered in the context of* P, i.e., in the context of the action synchronizations and hidings it is subject to in P. We denote by $C\langle P, k\rangle$ the behavior of C in the context of P, which is defined as follows: $C\langle P, k\rangle \xrightarrow{a,\lambda} C'\langle P', k\rangle$ iff $C \xrightarrow{a',\lambda'} C'$ is used in the derivation of $P \xrightarrow{a,\lambda} P'$ according to the rules in Table 1. We extend to behaviors of the form $C\langle P, k\rangle$ the notion of full instability.

For a computation of P to be reducible, the *same sequential components* must synchronize in *all steps* of the computation – with all the other sequential components staying idle – and each of those components must traverse *local states that are all fully unstable when considered in the context of* P. Since several reducible computations can depart from the same fully unstable local state (see the middle and rightmost parts of Fig. 1), in general we will have to handle *trees of reducible computations*, rather than individual reducible computations. Therefore, we further require that synchronizing local states of the involved sequential components have all the *same total exit rate when considered in the context of* P, which ensures that the sequential components at hand are synchronized with respect to an entire tree of reducible computations.

Definition 4.2. Let $n \in \mathbb{N}_{\geq 1}$, $P_1, P_2, \dots, P_{n+1} \in \mathbb{P}$ each having $\ell \in \mathbb{N}_{\geq 1}$ positions, and $\emptyset \neq K \subseteq \{1, \dots, \ell\}$. A computation c of length n from P_1 to P_{n+1} having the form $P_1 \xrightarrow{\tau,\lambda_1} P_2 \xrightarrow{\tau,\lambda_2} \dots \xrightarrow{\tau,\lambda_n} P_{n+1}$ is *locally reducible with respect to* K iff:

1. For all $i = 1, \dots, n$, transition $P_i \xrightarrow{\tau,\lambda_i} P_{i+1}$ is derived by applying rule PRE in Table 1 only to each sequential component occurring in position $k \in K$ of P_i.
2. For all $k \in K$, let $C_{k,1}, C_{k,2}, \dots, C_{k,n+1} \in \mathbb{P}_{\mathrm{seq}}$ with $C_{k,i}$ occurring in position k of P_i, $1 \leq i \leq n+1$, and $C_{k,1}\langle P_1, k\rangle \xrightarrow{\tau,\lambda_1} C_{k,2}\langle P_2, k\rangle \xrightarrow{\tau,\lambda_2} \dots \xrightarrow{\tau,\lambda_n} C_{k,n+1}\langle P_{n+1}, k\rangle$. Then for all $i = 1, \dots, n$:
   (a) $C_{k,i}\langle P_i, k\rangle$ is fully unstable.
   (b) $\mathit{rate}_{\mathrm{t}}(C_{k,i}\langle P_i, k\rangle) = \mathit{rate}_{\mathrm{t}}(C_{k',i}\langle P_i, k'\rangle)$ for all $k' \in K$.
3. Either $C_{k',n+1}\langle P_{n+1}, k'\rangle$ is not fully unstable for some $k' \in K$, or for all $k \in K$ it holds that $C_{k,n+1}\langle P_{n+1}, k\rangle = C_{k,j_k}\langle P_j, k\rangle$ for some $j_k = 1, \dots, n$ with at least one $k' \in K$ satisfying $C_{k',i}\langle P_i, k'\rangle \neq C_{k',j}\langle P_j, k'\rangle$ for all $i, j = 1, \dots, n$ such that $i \neq j$. ■

If locally reducible, each replica of the computation c above can be reduced to a replica of a single exponentially timed τ-transition tr that we consider equivalent to c, whose rate is obtained from the following generalization of the *probtime* measure by inverting its second factor:

$$\boxed{\mathit{probtime}_{\mathrm{g}}(c) \; = \; \left(\prod_{i=1}^{n} \frac{\lambda_i}{\mathit{rate}_{\mathrm{t}}(C_{k,i}\langle P_i, k\rangle)} \right) \cdot \left(\sum_{i=1}^{n} \frac{1}{\mathit{rate}_{\mathrm{t}}(C_{k,i}\langle P_i, k\rangle)} \right)}$$

Notice that, due to condition 2(b) in Def. 4.2, it is not important which specific position $k \in K$ is considered for computing $\mathit{probtime}_{\mathrm{g}}(c)$. Furthermore, the value of $\mathit{probtime}_{\mathrm{g}}(c)$ is the same for all replicas of c, as it does not depend on the total exit rates of the sequential components that stay idle, i.e., it does not depend on the contexts of the various replicas of c. When $\ell = 1$, $\mathit{probtime}_{\mathrm{g}}(c) = \mathit{probtime}(c)$.

Given $P \in \mathbb{P}$ having a computation locally reducible with respect to K, when calculating *pbtm* towards $D \subseteq \mathbb{P}$ we have to consider the entire tree of computations of P locally reducible with respect to K up to their final states, which are identified in condition 3 of Def. 4.2. Denoting by $\mathit{lrcomp}(P, D, t, K)$ the multiset of computations locally reducible with respect to K going from P to some final state in D whose expected duration is $t \in \mathbb{R}_{>0}$, what we need to compute is the following generalization of *pbtm* based on $\mathit{probtime}_{\mathrm{g}}$:

$$\boxed{\mathit{pbtm}_{\mathrm{g}}(P, D, K) \; = \bigcup_{t \in \mathbb{R}_{>0} \text{ s.t. } \mathit{lrcomp}(P,D,t,K) \neq \emptyset} \{\!| \sum_{c \in \mathit{lrcomp}(P,D,t,K)} \mathit{probtime}_{\mathrm{g}}(c) |\!\}}$$

4.2. Generalizing the Definition of $\approx_{\mathrm{MB}}$

Before introducing the generalized definition of weak Markovian bisimilarity, we need some further notation. In Def. 3.2, two checks are necessary: one for non-fully-unstable global states, based on *rate*, and one for fully unstable global states, based on *pbtm*. Here, the two checks have to be rephrased in terms of local states. As a consequence, the generalization of the second check will compare, according to $\mathit{pbtm}_{\mathrm{g}}$, locally reducible computations of different sets of synchronized sequential components. In contrast, the generalization of the first check will essentially take care of local states in which visible actions are enabled. To this purpose, a suitable generalization of *rate* will be defined.

In the following, $\mathcal{K}(P) = \mathcal{K}_{\mathrm{nlr}}(P) \cup \mathcal{K}_{\mathrm{lr}}(P)$ denotes the set of positions of $P \in \mathbb{P}$ – numbered 1 to ℓ_P – where $\mathcal{K}_{\mathrm{nlr}}(P)$ is the subset of positions k corresponding to all sequential components C of P such that $C\langle P, k\rangle$ has no locally reducible computations, while $\mathcal{K}_{\mathrm{lr}}(P)$ is the subset of positions k corresponding to all sequential components C of P such that $C\langle P, k\rangle$ has locally reducible computations. We will write $\mathcal{K}_{\mathrm{lr}}(P) = \mathcal{K}'_{\mathrm{lr}}(P) \uplus \mathcal{K}''_{\mathrm{lr}}(P)$ to denote the existence of two disjoint subsets $\mathcal{K}'_{\mathrm{lr}}(P)$ and $\mathcal{K}''_{\mathrm{lr}}(P)$ of $\mathcal{K}_{\mathrm{lr}}(P)$ whose union is $\mathcal{K}_{\mathrm{lr}}(P)$; in that case, we let $\mathcal{K}'_{\mathrm{nlr}}(P) = \mathcal{K}_{\mathrm{nlr}}(P) \cup \mathcal{K}'_{\mathrm{lr}}(P)$.

Moreover, given $a \in \mathit{Name}$, $D \subseteq \mathbb{P}$, and $K \subseteq \mathcal{K}(P)$, we define the following generalization of *rate*:

$$\boxed{\mathit{rate}_{\mathrm{g}}(P, a, D, K) \; = \; \sum_{k \in K} \mathit{rate}(C_k\langle P, k\rangle, a, D)}$$

where C_k is the sequential component occurring in position k of P. It is worth noting that, for $a \neq \tau$, $\mathit{rate}_{\mathrm{g}}(P, a, D, \mathcal{K}_{\mathrm{nlr}}(P)) = \mathit{rate}(P, a, D)$. Below, we again use $^+$ to denote the transitive closure operation.

Definition 4.3. A reflexive and symmetric relation $\mathcal{B}$ over $\mathbb{P}$ is a *generalized weak Markovian bisimulation* iff, whenever $(P_1, P_2) \in \mathcal{B}$, then $\mathcal{K}_{\mathrm{lr}}(P_1) = \mathcal{K}'_{\mathrm{lr}}(P_1) \uplus \mathcal{K}''_{\mathrm{lr}}(P_1)$ and $\mathcal{K}_{\mathrm{lr}}(P_2) = \mathcal{K}'_{\mathrm{lr}}(P_2) \uplus \mathcal{K}''_{\mathrm{lr}}(P_2)$ such that:

- For all action names $a \in \mathit{Name}$ and equivalence classes $D \in \mathbb{P}/\mathcal{B}^+$:
$$\mathit{rate}_{\mathrm{g}}(P_1, a, D, \mathcal{K}'_{\mathrm{nlr}}(P_1)) \; = \; \mathit{rate}_{\mathrm{g}}(P_2, a, D, \mathcal{K}'_{\mathrm{nlr}}(P_2))$$
- For each computation of P_1 locally reducible with respect to $K_1 \subseteq \mathcal{K}''_{\mathrm{lr}}(P_1)$ there exists a computation of P_2 locally reducible with respect to $K_2 \subseteq \mathcal{K}''_{\mathrm{lr}}(P_2)$ such that for all equivalence classes $D \in \mathbb{P}/\mathcal{B}^+$:
$$\mathit{pbtm}_{\mathrm{g}}(P_1, D, K_1) \; = \; \mathit{pbtm}_{\mathrm{g}}(P_2, D, K_2)$$
and vice versa.

We call *generalized weak Markovian bisimilarity*, denoted by $\approx_{\mathrm{MB,g}}$, the transitive closure of the largest generalized weak Markovian bisimulation. ■

We present below several examples showing the preservation of the identifications made possible by $\approx_{\rm MB}$, the recovery of congruence with respect to parallel composition, some motivations behind the factorization in Def. 4.3 of $\mathcal{K}_{\rm lr}(P_1)$ and $\mathcal{K}_{\rm lr}(P_2)$ that are related to the interleaving view of concurrency and the compositionality of hiding, and the need of transitive closure in Def. 4.3.

Example 4.4. The pairs of process terms examined in Exs. 3.3, 3.4, and 3.5 are still related by $\approx_{\rm MB,g}$. Each considered process term can be seen as a concurrent process term with a single sequential component not subject to hiding, so there is no difference between the behavior of the sequential component in the context of the overall process (global state view) and the behavior of the component in isolation (local state view). Moreover, the computations of those processes are locally reducible with respect to a singleton set of positions. Each of the first two processes has a single locally reducible computation (see the leftmost part of Fig. 1), while each of the other four processes has a tree with two locally reducible computations (see the middle and rightmost parts of Fig. 1). ■

Example 4.5. Let us reconsider the two process terms at the beginning of Ex. 4.1. Now we have that:

$$<a, \lambda>.<\tau, \mu>.<\tau, \gamma>.\underline{0} \approx_{\rm MB,g} <a, \lambda>.<\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0}$$

and also:

$$<a, \lambda>.<\tau, \mu>.<\tau, \gamma>.\underline{0} \parallel_{\emptyset} <a', \lambda'>.\underline{0} \approx_{\rm MB,g} <a, \lambda>.<\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0} \parallel_{\emptyset} <a', \lambda'>.\underline{0}$$

because it holds that:

$$<\tau, \mu>.<\tau, \gamma>.\underline{0} \parallel_{\emptyset} <a', \lambda'>.\underline{0} \approx_{\rm MB,g} <\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0} \parallel_{\emptyset} <a', \lambda'>.\underline{0}$$

In fact, for $a' \neq \tau$ each of the two process terms has a locally reducible computation with respect to $\{1\}$, i.e., originated by the leftmost sequential component. Each such computation has two replicas: the first one having context $\{<a', \lambda'>\}$ and final state $\underline{0} \parallel_{\emptyset} <a', \lambda'>.\underline{0}$, the second one having empty context and final state $\underline{0} \parallel_{\emptyset} \underline{0}$. It holds that:

$$pbtm_{\rm g}(<\tau, \mu>.<\tau, \gamma>.\underline{0} \parallel_{\emptyset} <a', \lambda'>.\underline{0}, D, \{1\}) \ = \ \{\!| \tfrac{1}{\mu} + \tfrac{1}{\gamma} |\!\}$$
$$pbtm_{\rm g}(<\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0} \parallel_{\emptyset} <a', \lambda'>.\underline{0}, D, \{1\}) \ = \ \{\!| \tfrac{\mu + \gamma}{\mu \cdot \gamma} |\!\}$$

for D containing the final state $\underline{0} \parallel_{\emptyset} <a', \lambda'>.\underline{0}$, as the way of calculating $probtime_{\rm g}$ and $pbtm_{\rm g}$ does not take the context of the replica into account.
For $a' = \tau$, each of the two process terms has an additional locally reducible computation, this time with respect to $\{2\}$, i.e., originated by the rightmost sequential component. Each such computation has two replicas of length 1 labeled with $<a', \lambda'>$. In this case, it holds that:

$$pbtm_{\rm g}(<\tau, \mu>.<\tau, \gamma>.\underline{0} \parallel_{\emptyset} <a', \lambda'>.\underline{0}, D, \{2\}) \ = \ \{\!| \tfrac{1}{\lambda'} |\!\}$$
$$pbtm_{\rm g}(<\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0} \parallel_{\emptyset} <a', \lambda'>.\underline{0}, D, \{2\}) \ = \ \{\!| \tfrac{1}{\lambda'} |\!\}$$

for D containing the two $\approx_{\rm MB,g}$-equivalent final states $<\tau, \mu>.<\tau, \gamma>.\underline{0} \parallel_{\emptyset} \underline{0}$ and $<\tau, \frac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0} \parallel_{\emptyset} \underline{0}$.
The two divergent process terms at the end of Ex. 4.1 are not related by $\approx_{\rm MB,g}$ because $\gamma_1 \neq \gamma_2$; hence, they no longer result in a disruption of compositionality when placed in the context of $\_ \parallel_{\emptyset} <a', \lambda'>.\underline{0}$. ■

Example 4.6. The factorization of $\mathcal{K}_{\rm lr}(P_1)$ and $\mathcal{K}_{\rm lr}(P_2)$ in Def. 4.3 allows processes with a different number of sequential components to be identified also in the presence of τ-actions.
Consider for instance the two $\approx_{\rm MB,g}$-equivalent process terms:

$$P_1 \ \equiv \ <a, \lambda>.<\tau, \mu>.\underline{0} + <\tau, \mu>.<a, \lambda>.\underline{0}$$
$$P_2 \ \equiv \ <a, \lambda>.\underline{0} \parallel_{\emptyset} <\tau, \mu>.\underline{0}$$

where the former ($\ell_{P_1} = 1$) is the interleaving view of the latter ($\ell_{P_2} = 2$), hence their underlying labeled multitransition systems are isomorphic. If the first check of Def. 4.3 were based on $\mathcal{K}_{\rm nlr}(P_1)$ and $\mathcal{K}_{\rm nlr}(P_2)$ alone, and the second one on the entire $\mathcal{K}_{\rm lr}(P_1)$ and $\mathcal{K}_{\rm lr}(P_2)$, then both checks would fail.
Observing that $\mathcal{K}_{\rm nlr}(P_1) = \{1\} = \mathcal{K}_{\rm nlr}(P_2)$, it would indeed hold that:

$$rate_{\rm g}(P_1, \tau, D, \mathcal{K}_{\rm nlr}(P_1)) \ = \ \mu \ \neq \ 0 \ = \ rate_{\rm g}(P_2, \tau, D, \mathcal{K}_{\rm nlr}(P_2))$$

for D containing $<a, \lambda>.\underline{0}$. Moreover, while P_2 has a computation locally reducible with respect to $\{2\}$, it turns out that P_1 has no locally reducible computations at all, which would violate the second check.
The flexibility inherent to Def. 4.3 permits to factorize $\mathcal{K}_{\rm lr}(P_2)$ as $\{2\} \uplus \emptyset$ so that $\mathcal{K}'_{\rm nlr}(P_2) = \{1, 2\}$ and:

$$rate_{\rm g}(P_1, \tau, D, \mathcal{K}'_{\rm nlr}(P_1)) \ = \ \mu \ = \ rate_{\rm g}(P_2, \tau, D, \mathcal{K}'_{\rm nlr}(P_2))$$

for D containing the two $\approx_{\rm MB,g}$-equivalent process terms $<a, \lambda>.\underline{0} \parallel_{\emptyset} \underline{0}$ and $<a, \lambda>.\underline{0}$. On the other hand, $\mathcal{K}''_{\rm lr}(P_1) = \emptyset = \mathcal{K}''_{\rm lr}(P_2)$ and hence the second check is trivially passed.

We further mention that this flexibility is also necessary to achieve compositionality with respect to hiding. For example, Q_1 and Q_2, respectively obtained from P_1 and P_2 by replacing every occurrence of τ with b, are $\approx_{\mathrm{MB,g}}$-equivalent but, with the unflexible checks, $Q_1/\{b\}$ and $Q_2/\{b\}$ would no longer be so, as they are respectively isomorphic to P_1 and P_2. ■

Example 4.7. A generalized weak Markovian bisimulation is not necessarily transitive, which is the reason for explicitly using transitive closure in Def. 4.3. In addition to P_1 and P_2 defined in Ex. 4.6, consider for instance the following process term:

$$P_3 \;\equiv\; <a, \lambda>.\underline{0} \,\|_{\emptyset}\, <\tau, 2 \cdot \mu>.<\tau, 2 \cdot \mu>.\underline{0}$$

It is easy to find a generalized weak Markovian bisimulation $\mathcal{B}_{1,2}$ containing (P_1, P_2), and another one $\mathcal{B}_{2,3}$ containing (P_2, P_3), but there is no generalized weak Markovian bisimulation containing (P_1, P_3). In order for $P_1 \approx_{\mathrm{MB,g}} P_3$ to hold, the transitive closure of $\mathcal{B}_{1,2} \cup \mathcal{B}_{2,3}$ has to be taken into account. ■

We conclude by showing that $\approx_{\mathrm{MB,g}}$ is a *conservative extension* of $\approx_{\mathrm{MB}}$ only for process terms that have *no cycles of exponentially timed τ-transitions*. The reason of this limitation is that $\approx_{\mathrm{MB,g}}$ imposes checks on those cycles that are not always performed by $\approx_{\mathrm{MB}}$, like, e.g., in the case of the two divergent process terms $\mathrm{rec}\, X : <\tau, \gamma_1>.X$ and $\mathrm{rec}\, X : <\tau, \gamma_2>.X$ such that $\gamma_1 \neq \gamma_2$. Moreover, $\approx_{\mathrm{MB,g}}$ and $\approx_{\mathrm{MB}}$ coincide over *non-divergent sequential processes*, because in that case the notion of locally reducible computation of Def. 4.2 boils down to the notion of reducible computation of Def. 3.1.

Proposition 4.8. Let $P_1, P_2 \in \mathbb{P}$ be not divergent. Then:

$$P_1 \approx_{\mathrm{MB}} P_2 \;\Longrightarrow\; P_1 \approx_{\mathrm{MB,g}} P_2$$ ■

Corollary 4.9. Let $P_1, P_2 \in \mathbb{P}_{\mathrm{seq}}$ be not divergent. Then:

$$P_1 \approx_{\mathrm{MB,g}} P_2 \;\Longleftrightarrow\; P_1 \approx_{\mathrm{MB}} P_2$$ ■

4.3. Congruence Property

The investigation of the compositionality of $\approx_{\mathrm{MB,g}}$ with respect to MPC operators leads to results analogous to those for $\approx_{\mathrm{MB}}$, with the additional achievement of congruence with respect to parallel composition.

Proposition 4.10. Let $P_1, P_2 \in \mathbb{P}$. If $P_1 \approx_{\mathrm{MB,g}} P_2$, then:

1. $<a, \lambda>.P_1 \approx_{\mathrm{MB,g}} <a, \lambda>.P_2$ for all $<a, \lambda> \in \mathit{Act}$ (when $P_1, P_2 \in \mathbb{P}_{\mathrm{seq}}$).
2. $P_1/H \approx_{\mathrm{MB,g}} P_2/H$ for all $H \subseteq \mathit{Name}_{\mathrm{v}}$.
3. $P_1 \,\|_S\, P \approx_{\mathrm{MB,g}} P_2 \,\|_S\, P$ and $P \,\|_S\, P_1 \approx_{\mathrm{MB,g}} P \,\|_S\, P_2$ for all $S \subseteq \mathit{Name}_{\mathrm{v}}$ and $P \in \mathbb{P}$. ■

Like $\approx_{\mathrm{MB}}$, the relation $\approx_{\mathrm{MB,g}}$ is not a congruence with respect to the alternative composition operator because, for instance, it holds that:

$$<\tau, \mu>.<\tau, \gamma>.\underline{0} \;\approx_{\mathrm{MB,g}}\; <\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0}$$

but:

$$<\tau, \mu>.<\tau, \gamma>.\underline{0} + <a, \lambda>.\underline{0} \;\not\approx_{\mathrm{MB,g}}\; <\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0} + <a, \lambda>.\underline{0}$$

In fact, if it were $a \neq \tau$, then we would have:

$$\mathit{rate}_{\mathrm{g}}(<\tau, \mu>.<\tau, \gamma>.\underline{0} + <a, \lambda>.\underline{0},\ \tau,\ [\underline{0}]_{\approx_{\mathrm{MB,g}}}, \{1\}) \;=\; 0$$
$$\mathit{rate}_{\mathrm{g}}(<\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0} + <a, \lambda>.\underline{0},\ \tau,\ [\underline{0}]_{\approx_{\mathrm{MB,g}}}, \{1\}) \;=\; \tfrac{\mu \cdot \gamma}{\mu + \gamma}$$

otherwise for $a = \tau$ each of the two process terms would have a single tree of computations locally reducible with respect to $\{1\}$ with final state $\underline{0}$ and we would have:

$$\mathit{pbtm}_{\mathrm{g}}(<\tau, \mu>.<\tau, \gamma>.\underline{0} + <a, \lambda>.\underline{0},\ [\underline{0}]_{\approx_{\mathrm{MB,g}}},\ \{1\}) \;=\; \{\!\!\{\ \tfrac{\mu}{\mu + \lambda} \cdot \left(\tfrac{1}{\mu + \lambda} + \tfrac{1}{\gamma} \right), \tfrac{\lambda}{\mu + \lambda} \cdot \tfrac{1}{\mu + \lambda}\ \}\!\!\}$$
$$\mathit{pbtm}_{\mathrm{g}}(<\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0} + <a, \lambda>.\underline{0},\ [\underline{0}]_{\approx_{\mathrm{MB,g}}},\ \{1\}) \;=\; \{\!\!\{\ \tfrac{1}{\frac{\mu \cdot \gamma}{\mu + \gamma} + \lambda}\ \}\!\!\}$$

The congruence violation can be solved as in Sect. 3.2, with the resulting equivalence $\simeq_{\mathrm{MB,g}}$ being the coarsest congruence – with respect to alternative composition – contained in $\approx_{\mathrm{MB,g}}$.

Definition 4.11. Let $P_1, P_2 \in \mathbb{P}$. We say that P_1 is *generalized weakly Markovian bisimulation congruent* to P_2, written $P_1 \simeq_{\mathrm{MB,g}} P_2$, iff for all action names $a \in \mathit{Name}$ and equivalence classes $D \in \mathbb{P}/\approx_{\mathrm{MB,g}}$:

$$\mathit{rate}(P_1, a, D) \ = \ \mathit{rate}(P_2, a, D) \qquad \blacksquare$$

Proposition 4.12. $\sim_{\mathrm{MB}} \subset \simeq_{\mathrm{MB,g}} \subset \approx_{\mathrm{MB,g}}$, with $\simeq_{\mathrm{MB,g}} = \approx_{\mathrm{MB,g}}$ over the set of process terms of $\mathbb{P}$ that have no locally reducible computations. ■

Proposition 4.13. Let $P_1, P_2 \in \mathbb{P}_{\mathrm{seq}}$ and $<a, \lambda> \in \mathit{Act}$. Then $<a, \lambda>.P_1 \simeq_{\mathrm{MB,g}} <a, \lambda>.P_2$ iff $P_1 \approx_{\mathrm{MB,g}} P_2$. ■

Theorem 4.14. Let $P_1, P_2 \in \mathbb{P}$. If $P_1 \simeq_{\mathrm{MB,g}} P_2$, then:

1. $<a, \lambda>.P_1 \simeq_{\mathrm{MB,g}} <a, \lambda>.P_2$ for all $<a, \lambda> \in \mathit{Act}$ (when $P_1, P_2 \in \mathbb{P}_{\mathrm{seq}}$).
2. $P_1 + P \simeq_{\mathrm{MB,g}} P_2 + P$ and $P + P_1 \simeq_{\mathrm{MB,g}} P + P_2$ for all $P \in \mathbb{P}_{\mathrm{seq}}$ (when $P_1, P_2 \in \mathbb{P}_{\mathrm{seq}}$).
3. $P_1 / H \simeq_{\mathrm{MB,g}} P_2 / H$ for all $H \subseteq \mathit{Name}_{\mathrm{v}}$.
4. $P_1 \parallel_S P \simeq_{\mathrm{MB,g}} P_2 \parallel_S P$ and $P \parallel_S P_1 \simeq_{\mathrm{MB,g}} P \parallel_S P_2$ for all $S \subseteq \mathit{Name}_{\mathrm{v}}$ and $P \in \mathbb{P}$. ■

Theorem 4.15. Let $P_1, P_2 \in \mathbb{P}_{\mathrm{seq}}$. Then $P_1 \simeq_{\mathrm{MB,g}} P_2$ iff $P_1 + P \approx_{\mathrm{MB,g}} P_2 + P$ for all $P \in \mathbb{P}_{\mathrm{seq}}$. ■

Definition 4.16. Let $P_1, P_2 \in \mathcal{PL}$ be process terms containing free occurrences of $k \in \mathbb{N}$ process variables $X_1, \dots, X_k \in \mathit{Var}$ at most. We define $P_1 \simeq_{\mathrm{MB,g}} P_2$ iff $P_1\{Q_i \hookrightarrow X_i \mid 1 \leq i \leq k\} \simeq_{\mathrm{MB,g}} P_2\{Q_i \hookrightarrow X_i \mid 1 \leq i \leq k\}$ for all $Q_1, \dots, Q_k \in \mathcal{PL}$ containing no free occurrences of process variables. ■

Definition 4.17. A binary relation $\mathcal{B}$ over $\mathbb{P}$ is a *generalized weak Markovian bisimulation up to* $\approx_{\mathrm{MB,g}}$ iff, whenever $(P_1, P_2) \in \mathcal{B}$, then $\mathcal{K}_{\mathrm{lr}}(P_1) = \mathcal{K}'_{\mathrm{lr}}(P_1) \uplus \mathcal{K}''_{\mathrm{lr}}(P_1)$ and $\mathcal{K}_{\mathrm{lr}}(P_2) = \mathcal{K}'_{\mathrm{lr}}(P_2) \uplus \mathcal{K}''_{\mathrm{lr}}(P_2)$ such that:

- For all action names $a \in \mathit{Name}$ and equivalence classes $D \in \mathbb{P}/(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+$:
$$\mathit{rate}_{\mathrm{g}}(P_1, a, D, \mathcal{K}'_{\mathrm{nlr}}(P_1)) \ = \ \mathit{rate}_{\mathrm{g}}(P_2, a, D, \mathcal{K}'_{\mathrm{nlr}}(P_2))$$
- For each computation of P_1 locally reducible with respect to $K_1 \subseteq \mathcal{K}''_{\mathrm{lr}}(P_1)$ there exists a computation of P_2 locally reducible with respect to $K_2 \subseteq \mathcal{K}''_{\mathrm{lr}}(P_2)$ such that for all equivalence classes $D \in \mathbb{P}/(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+$:
$$\mathit{pbtm}_{\mathrm{g}}(P_1, D, K_1) \ = \ \mathit{pbtm}_{\mathrm{g}}(P_2, D, K_2)$$
and vice versa. ■

Proposition 4.18. Let $\mathcal{B}$ be a relation over $\mathbb{P}$. If $\mathcal{B}$ is a generalized weak Markovian bisimulation up to $\approx_{\mathrm{MB,g}}$, then $(P_1, P_2) \in \mathcal{B}$ implies $P_1 \approx_{\mathrm{MB,g}} P_2$ for all $P_1, P_2 \in \mathbb{P}$. Moreover $(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+ = \approx_{\mathrm{MB,g}}$. ■

Theorem 4.19. Let $P_1, P_2 \in \mathcal{PL}$ be sequential process terms containing free occurrences of $k \in \mathbb{N}$ process variables $X_1, \dots, X_k \in \mathit{Var}$ at most. Whenever $P_1 \simeq_{\mathrm{MB,g}} P_2$, then:

$$\mathrm{rec}\, X_1 : \dots : \mathrm{rec}\, X_k : P_1 \ \simeq_{\mathrm{MB,g}} \ \mathrm{rec}\, X_1 : \dots : \mathrm{rec}\, X_k : P_2 \qquad \blacksquare$$

4.4. Exactness at Stationary State

The CTMC-level pseudo-aggregation induced by $\approx_{\mathrm{MB,g}}$ and $\simeq_{\mathrm{MB,g}}$ – which we call GW-lumpability – shares with the one induced by $\approx_{\mathrm{MB}}$ and $\simeq_{\mathrm{MB}}$ – called W-lumpability in Sect. 3.4 – the characteristic of viewing a sequence of exponentially timed τ-transitions as equivalent to an individual exponentially timed τ-transition having the same expected duration and the same execution probability. However, in the case of GW-lumpability, due to the idea of context embodied in the notion of locally reducible computation, the equivalence holds when the transitions are considered *locally* to the processes originating them. Thus, the main difference between GW-lumpability and W-lumpability is that the former may aggregate states also in the presence of concurrent processes performing other transitions, while the latter cannot.

The rewriting rule describing GW-lumpability is presented in Fig. 6. The sum Λ of the rates of the other transitions departing from the considered states is the overall rate of the context. The states s_1 to $s_{|I|}$ reachable in one step not only from s should be duplicated before the aggregation takes place. This implies in particular that aggregations along different replicas of the same locally reducible computation can be done separately. The two-step structural definition of GW-lumpability is a generalization of Def. 3.23.

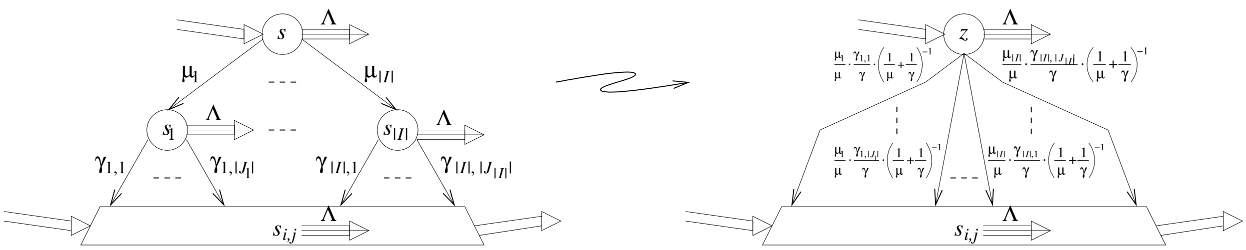

Figure 6: Effect of GW-lumpability ($\mu = \sum_{i \in I} \mu_i$, $\gamma = \sum_{j \in J_i} \gamma_{i,j}$ for all $i \in I$)

Definition 4.20. Let (S, R) be a CTMC. An equivalence relation $\mathcal{R} \subseteq S \times S$ is a *GW-lumping* iff, whenever $(s_1, s_2) \in \mathcal{R}$, then one of the following three conditions holds:

- For all equivalence classes $D \in S/\mathcal{R}$:
$$R(s_1, D) \;=\; R(s_2, D)$$
- For one of s_1 and s_2, which we denote by s, there exist $\mu, \gamma \in \mathbb{R}_{>0}$ and $\Lambda \in \mathbb{R}_{\geq 0}$ such that:
  - $E(s) = \mu + \Lambda$,
  - $E(s') = \gamma + \Lambda$ for all $s' \in S$ such that $R(s, s') > 0$,
  - $\{s'' \in S \mid R(s'', s') > 0\} = \{s\}$ for all $s' \in S$ such that $R(s, s') > 0$,

  while for the other one of s_1 and s_2, which we denote by z, it holds that:
  - $E(z) = \left(\frac{1}{\mu} + \frac{1}{\gamma}\right)^{-1} + \Lambda$.

  Moreover, for all equivalence classes $D \in S/\mathcal{R}$:
$$\frac{R(z,D)}{E(z)-\Lambda} \;=\; \sum_{D' \in S/\mathcal{R}} \left(\frac{R(s,D')}{E(s)-\Lambda} \cdot \sum_{s' \in D' \text{ s.t. } R(s,s')>0} \frac{R(s',D)}{E(s')-\Lambda} \right)$$
- For each $i = 1, 2$ there exist $\mu_i, \gamma_i \in \mathbb{R}_{>0}$ and $\Lambda \in \mathbb{R}_{\geq 0}$ such that:
  - $E(s_i) = \mu_i + \Lambda$.
  - $E(s') = \gamma_i + \Lambda$ for all $s' \in S$ such that $R(s_i, s') > 0$.
  - $\{s'' \in S \mid R(s'', s') > 0\} = \{s_i\}$ for all $s' \in S$ such that $R(s_i, s') > 0$.

  Moreover:
$$\left(\frac{1}{\mu_1} + \frac{1}{\gamma_1}\right)^{-1} \;=\; \left(\frac{1}{\mu_2} + \frac{1}{\gamma_2}\right)^{-1}$$
  and for all equivalence classes $D \in S/\mathcal{R}$:
$$\sum_{D' \in S/\mathcal{R}} \left(\frac{R(s_1,D')}{E(s_1)-\Lambda} \cdot \sum_{s' \in D' \text{ s.t. } R(s_1,s')>0} \frac{R(s',D)}{E(s')-\Lambda} \right) \;=\; \sum_{D' \in S/\mathcal{R}} \left(\frac{R(s_2,D')}{E(s_2)-\Lambda} \cdot \sum_{s' \in D' \text{ s.t. } R(s_2,s')>0} \frac{R(s',D)}{E(s')-\Lambda} \right)$$

We say that $s_1, s_2 \in S$ are *GW-lumpable* iff $(s_1, s_2) \in \mathcal{R}$ for some GW-lumping $\mathcal{R}$. ■

As seen at the beginning of Sect. 4.1, while W-lumpability preserves transient properties expressed in terms of the mean time to certain events, this is no longer the case with GW-lumpability. As far as stationary-state reward-based performance measures are concerned, it turns out that they are *conditionally* preserved by GW-lumpability. To be precise, stationary-state exactness of GW-lumpability holds as long as we confine ourselves to processes in which only synchronizations of a certain kind can take place before the beginning of locally reducible computations. Similar to W-lumpability, the states traversed by a replica of a locally reducible computation should be given equal cumulative rewards, and the transitions belonging to the replica should be given appropriate instantaneous rewards (e.g., null rewards).

The proof of *conditional* stationary-state exactness for GW-lumpability cannot follow the same path as the proof of *full* stationary-state exactness for W-lumpability of Thm. 3.24. The reason is that, when the overall rate Λ of the context is greater than zero, the global balance equations of the original CTMC cannot be transformed into a form equivalent to the global balance equations of the aggregated CTMC. For GW-lumpability, we follow instead the scheme of the proof of an analogous result for weak isomorphism [22], by viewing the considered CTMC models as generalized semi-Markov processes and then exploiting the insensitivity results for the latter models.

A *generalized semi-Markov process (GSMP)* is a stochastic process in which there are elements subject to birth and death when moving across states. The probability of each transition depends on the source state, the target state, and the element that completes its lifetime during the transition; all interrupted elements record their residual lifetimes. Some elements have an exponentially distributed lifetime, whilst the lifetime of the others is generally distributed. Such a distinction is important for those elements whose lifetime spans over several consecutive states. Indeed, only exponential distributions are memoryless and hence allow elements to be restarted instead of being resumed from one state to another, because in that case the distributions of residual lifetimes coincide with the distributions of the corresponding initial lifetimes.

In our behavioral setting, elements represent sets of actions enabled by local states of sequential components, so that the original and aggregated CTMC models can be regarded as GSMP models in which all the elements have exponentially distributed lifetimes. When proving conditional stationary-state exactness for GW-lumpability, we will consider an intermediate GSMP having the same state space as the aggregated one, whose transitions corresponding to locally reducible computations in the original GSMP are representative of elements whose lifetime is hypoexponentially distributed. With regard to Fig. 6, the transitions of the state z' of the intermediate GSMP corresponding to the transitions of state z that are explicitly depicted in the figure, are associated with an element following a hypoexponential distribution with two stages having rate μ and γ, respectively. The death of this element causes z' to reach state $s'_{i,j}$ of the intermediate GSMP corresponding to state $s_{i,j}$ with probability $\frac{\mu_i}{\mu} \cdot \frac{\gamma_{i,j}}{\gamma}$.

According to Matthes' theorem [26], a GSMP is *insensitive* to the elements whose lifetimes are generally distributed – in the sense that the stationary state probability distribution of the GSMP does not change in the case that those general distributions are replaced by any other distributions with the same means – iff, when the lifetimes of all the generally distributed elements are assumed to be exponentially distributed, the flux out of each state due to the death of one of those elements is equal to the flux into the same state due to the birth of that element. The condition about flux equality yields the so called insensitivity balance equations (see, e.g., [5]). These equations are the ones that will be considered (especially for the intermediate GSMP) in the proof of stationary-state exactness for GW-lumpability in place of the global balance equations.

Following [26], in order to ensure the insensitivity of a GSMP, a state transition (i) cannot be due to the simultaneous death of two generally distributed elements and (ii) cannot cause the simultaneous birth of two generally distributed elements, with all the other generally distributed elements retaining their residual lifetimes. In our setting, elements must belong to different sequential components and therefore they carry over their residual lifetimes when they are interrupted. Moreover, two elements cannot die simultaneously, because the sequential components to which they belong are subject to a race and the distributions of their lifetimes are continuous, hence the probability of terminating in the same instant is zero. In contrast, the simultaneous birth of two (generally distributed) elements is possible (in the intermediate GSMP) after a synchronization between the two sequential components to which the elements belong.

Similar to the case of two independent sequential components (i.e., two components that never synchronize with each other), two fully synchronized sequential components (i.e., two components that synchronize on all the actions enabled by one of them, with the other possibly performing additional actions autonomously) cannot violate insensitivity because they can be viewed as a single sequential component. Likewise, rephrasing [19], two partially synchronized components can be considered as *joining* together after a synchronization if at any subsequent stage of their lifetimes (i) both of them are interrupted, (ii) only one

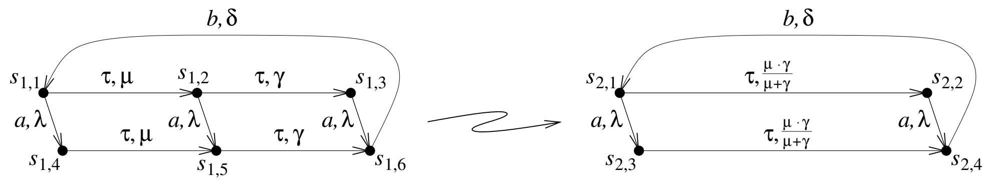

Figure 7: GW-lumpability is not exact in this case due to a non-joining synchronization

of them advances while the other is interrupted, or (iii) neither is interrupted and they synchronize again with each other or only one of them advances autonomously. In other words, there is no subsequent stage at which they can *both* proceed *autonomously.* The previous considerations lead to the following definition, which is then used to express a synchronization-related constraint in the theorem below.

Definition 4.21. Let C_1 and C_2 be two sequential components occurring in positions k_1 and k_2 of $P \in \mathbb{P}$, respectively. Let s be a (global) state of $[\![P]\!]$ and denote by $C_1^s\langle P, k_1\rangle$ and $C_2^s\langle P, k_2\rangle$ the local states of C_1 and C_2 in s, respectively. We say that a synchronization taking place at state s between $C_1^s\langle P, k_1\rangle$ and $C_2^s\langle P, k_2\rangle$ *joins* C_1 and C_2 from s on if in every state s' reachable in one or more steps from s it holds that at most one of $C_1^{s'}\langle P, k_1\rangle$ and $C_2^{s'}\langle P, k_2\rangle$ can advance without synchronizing with the other, otherwise we say that the synchronization at s is *non-joining.* ■

Theorem 4.22. GW-lumpability is exact at stationary state over every CTMC underlying a process $P \in \mathbb{P}$ such that, for each locally reducible computation in $[\![P]\!]$, the initial state of the computation is not reachable in one or more steps from a state in which a non-joining synchronization takes place. ■

Example 4.23. The constraint on synchronizations is explicitly used in the proof of Thm. 4.22. However, to illustrate better the need of it with an example, consider the following two process terms:

$$P_1 \;\equiv\; (\mathrm{rec}\, X : <\tau, \mu>.<\tau, \gamma>.<b, \delta>.X) \parallel_{\{b\}} (\mathrm{rec}\, Y : <a, \lambda>.<b, \delta>.Y)$$
$$P_2 \;\equiv\; (\mathrm{rec}\, X : <\tau, \tfrac{\mu\cdot\gamma}{\mu+\gamma}>.<b, \delta>.X) \parallel_{\{b\}} (\mathrm{rec}\, Y : <a, \lambda>.<b, \delta>.Y)$$

Observe that $P_1 \approx_{\mathrm{MB,g}} P_2$, where $[\![P_1]\!]$ and $[\![P_2]\!]$ are depicted in Fig. 7. They respectively extend the two labeled multitransition systems in Fig. 4 with an additional transition labeled with $<b, \delta>$ (assume $\delta \otimes \delta = \delta$ for the rate synchronization operator $\otimes$ in Table 1) going from the final state ($s_{1,6}$ and $s_{2,4}$, respectively) back to the initial one ($s_{1,1}$ and $s_{2,1}$, respectively). In the case that $\mu = \gamma = \lambda = \delta = 1$, it turns out that the stationary state probability distribution for $[\![P_1]\!]$ is as follows:

$$\pi[s_{1,1}] = \tfrac{2}{13} \qquad \pi[s_{1,2}] = \tfrac{1}{13} \qquad \pi[s_{1,3}] = \tfrac{1}{13}$$
$$\pi[s_{1,4}] = \tfrac{2}{13} \qquad \pi[s_{1,5}] = \tfrac{3}{13} \qquad \pi[s_{1,6}] = \tfrac{4}{13}$$

whereas the stationary state probability distribution for $[\![P_2]\!]$ is as follows:

$$\pi[s_{2,1}] = \tfrac{2}{10} \qquad \pi[s_{2,2}] = \tfrac{1}{10}$$
$$\pi[s_{2,3}] = \tfrac{4}{10} \qquad \pi[s_{2,4}] = \tfrac{3}{10}$$

Thus, the CTMC underlying $[\![P_2]\!]$ is not an exact pseudo-aggregation of the one underlying $[\![P_1]\!]$ because:

$$\pi[s_{1,1}] + \pi[s_{1,2}] \neq \pi[s_{2,1}] \qquad \pi[s_{1,3}] \neq \pi[s_{2,2}]$$
$$\pi[s_{1,4}] + \pi[s_{1,5}] \neq \pi[s_{2,3}] \qquad \pi[s_{1,6}] \neq \pi[s_{2,4}]$$

As can be noted, the transition in $[\![P_1]\!]$ labeled with $<b, \delta>$ is a non-joining synchronization, because the two sequential components can proceed autonomously in the subsequent states $s_{1,1}$ and $s_{1,2}$; hence, Thm. 4.22 does not apply. It is worth observing that the lack of stationary-state exactness is not caused by the fact that the target state of the non-joining synchronization transition is the initial state of a locally reducible computation. The problem is the existence of the non-joining synchronization itself. Suppose that the leftmost sequential component of P_1 is $\mathrm{rec}\, X : <\tau, \mu>.<\tau, \rho_1>.<\tau, \rho_2>.<\tau, \gamma>.<b, \delta>.X$ and that the leftmost sequential component of P_2 is $\mathrm{rec}\, X : <\tau, \mu>.<\tau, \frac{\rho_1\cdot\rho_2}{\rho_1+\rho_2}>.<\tau, \gamma>.<b, \delta>.X$, so to avoid any direct interference with the non-joining synchronization. Not even in this case the pseudo-aggregation is exact.

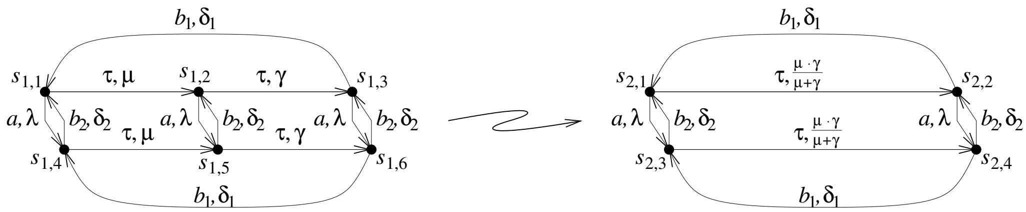

Figure 8: GW-lumpability is exact in this case due to the absence of synchronizations

In contrast, the result applies if we consider the following synchronization-free variant of the two process terms above:

$$P_3 \equiv (\mathrm{rec}\, X : <\tau, \mu>.<\tau, \gamma>.<b_1, \delta_1>.X) \parallel_\emptyset (\mathrm{rec}\, Y : <a, \lambda>.<b_2, \delta_2>.Y)$$
$$P_4 \equiv (\mathrm{rec}\, X : <\tau, \tfrac{\mu \cdot \gamma}{\mu+\gamma}>.<b_1, \delta_1>.X) \parallel_\emptyset (\mathrm{rec}\, Y : <a, \lambda>.<b_2, \delta_2>.Y)$$

where $[\![P_3]\!]$ and $[\![P_4]\!]$ are depicted in Fig. 8. For $\mu = \gamma = \lambda = \delta_1 = \delta_2 = 1$, we have that the stationary state probability distribution for $[\![P_3]\!]$ is:

$$\pi[s_{3,1}] = \tfrac{1}{6} \qquad \pi[s_{3,2}] = \tfrac{1}{6} \qquad \pi[s_{3,3}] = \tfrac{1}{6}$$
$$\pi[s_{3,4}] = \tfrac{1}{6} \qquad \pi[s_{3,5}] = \tfrac{1}{6} \qquad \pi[s_{3,6}] = \tfrac{1}{6}$$

and the stationary state probability distribution for $[\![P_4]\!]$ is:

$$\pi[s_{4,1}] = \tfrac{2}{6} \qquad \pi[s_{4,2}] = \tfrac{1}{6}$$
$$\pi[s_{4,3}] = \tfrac{2}{6} \qquad \pi[s_{4,4}] = \tfrac{1}{6}$$

hence the CTMC underlying $[\![P_4]\!]$ is an exact pseudo-aggregation of the one underlying $[\![P_3]\!]$ because:

$$\pi[s_{3,1}] + \pi[s_{3,2}] = \pi[s_{4,1}] \qquad \pi[s_{3,3}] = \pi[s_{4,2}]$$
$$\pi[s_{3,4}] + \pi[s_{3,5}] = \pi[s_{4,3}] \qquad \pi[s_{3,6}] = \pi[s_{4,4}]$$

■

4.5. Coarser Generalized Weak Markovian Bisimilarities

Similar to Sect. 3.5, instead of using $pbtm_{\mathrm{g}}(P, D, K)$ defined as the t-indexed multiset of sums of $probtime_{\mathrm{g}}$ measures over all locally reducible computations with respect to K from P to some final state in D *having the same expected duration* t, we could resort to $pbtm'_{\mathrm{g}}(P, D, K)$ computed by adding up the $probtime_{\mathrm{g}}$ measures of all the computations above *without considering their expected durations.* This results in a coarser generalized weak Markovian bisimulation equivalence $\approx'_{\mathrm{MB,g}}$ and the corresponding $\simeq'_{\mathrm{MB,g}}$, which respectively enjoy the same properties as $\approx_{\mathrm{MB,g}}$ and $\simeq_{\mathrm{MB,g}}$.

5. Equivalence Checking and State Space Reduction

We now address some algorithmic issues related to the verification of the various weak Markovian bisimilarities that we have introduced. In Sect. 5.1, we exhibit a decision procedure for $\approx_{\mathrm{MB}}$ and $\simeq_{\mathrm{MB}}$, which works in polynomial time over finite-state processes having no cycles of exponentially timed internal transitions. This procedure can be exploited for compositional state space reduction with respect to $\approx_{\mathrm{MB,g}}$ and $\simeq_{\mathrm{MB,g}}$, as will be illustrated in Sect. 5.2.

5.1. Polynomial-Time Decidability of $\approx_{\mathrm{MB}}$ and $\simeq_{\mathrm{MB}}$

To check whether $P_1 \approx_{\mathrm{MB}} P_2$ or $P_1 \simeq_{\mathrm{MB}} P_2$ for two finite-state processes $P_1, P_2 \in \mathbb{P}$, similar to other bisimulation equivalences we employ a partition refinement algorithm based on [30] that works as follows:

- Start with a partition containing an equivalence class for all the non-fully-unstable states of $[\![P_1]\!]$ and $[\![P_2]\!]$, together with another equivalence class for all the fully unstable states of $[\![P_1]\!]$ and $[\![P_2]\!]$.
- Refine the partition until a fixed point is reached, by applying the *rate*-based equality check for splitting the classes of non-fully-unstable states and the *pbtm*-based equality check for splitting the classes of fully unstable states.

- For $\approx_{\mathrm{MB}}$, return yes/no depending on whether P_1 and P_2 belong to the same equivalence class.
- For $\simeq_{\mathrm{MB}}$, return yes/no depending on whether P_1 and P_2 belong to the same equivalence class and satisfy the *rate*-based equality check with respect to all action names and equivalence classes.

In the case that the algorithm provides a positive answer, if the state-transition model underlying one of the two process terms is known to be smaller than the other, then the former process term can be exploited to verify functional properties or assess performance measures more quickly also for the latter process term. Likewise, the same algorithm can be employed to minimize the state-transition model underlying a single process term, so to make analysis faster.

Unlike weak bisimulation equivalences for nondeterministic processes and probabilistic processes – which can be decided in polynomial time for all pairs of finite-state processes with analogous partition refinement algorithms [24, 6, 4] – the above algorithm executes in polynomial time only when $[\![P_1]\!]$ and $[\![P_2]\!]$ have no cycles of exponentially timed internal transitions. In fact, while cycles of nondeterministic internal transitions are unimportant from a quantitative viewpoint, and cycles of probabilistic internal transitions can be left in the long run with probability 1 (if admitting a way out) or 0 (if connecting an absorbing set of states), cycles of exponentially timed internal transitions cause time to progress. In particular, their presence causes *pbtm* multisets to become infinite.

Example 5.1. Consider $P \equiv {<}\tau, \mu{>}.\, \mathrm{rec}\, X : ({<}\tau, \delta{>}.X + {<}\tau, \gamma{>}.Q)$ where $Q \in \mathbb{P}_{\mathrm{nfu}}$. Due to the presence in $[\![P]\!]$ of the exponentially timed internal selfloop labeled with ${<}\tau, \delta{>}$, we have that $\mathit{pbtm}(P, [Q]_{\approx_{\mathrm{MB}}})$ contains infinitely many *probtime* values of the form $(\frac{\mu}{\mu} \cdot (\frac{\delta}{\delta+\gamma})^n \cdot \frac{\gamma}{\delta+\gamma}) \cdot (\frac{1}{\mu} + (n+1) \cdot \frac{1}{\delta+\gamma})$ where $n \in \mathbb{N}$. ■

5.2. An Example of Compositional State Space Reduction via $\simeq_{\mathrm{MB,g}}$

Since the relation $\simeq_{\mathrm{MB,g}}$ is a congruence with respect to all the operators of MPC (see Thms. 4.14 and 4.19), it can be exploited for compositional state space reduction. Due to the fact that $\simeq_{\mathrm{MB,g}}$ coincides with $\simeq_{\mathrm{MB}}$ over non-divergent sequential components (see Prop. 4.8), the idea is to apply the partition refinement algorithm for $\simeq_{\mathrm{MB}}$ to each sequential component of the process under consideration, after hiding all the appropriate actions. The model resulting from the parallel composition of the minimized sequential components will preserve stationary-state reward-based performance measures if no locally reducible computation of the original model is directly or indirectly preceded by a non-joining synchronization (see Thm. 4.22).

To illustrate $\simeq_{\mathrm{MB,g}}$ at work, we consider the dining philosophers problem. Suppose that there are $n \geq 2$ philosophers sitting at a round table. After thinking for a while, everyone of them needs to eat some rice. Before eating, the philosopher has to get both the chopstick on the left and the chopstick on the right, with each chopstick being shared by two neighboring philosophers and usable by only one of them at a time. After eating, the philosopher has to put both chopsticks on the table, and then starts thinking again. This problem involves sequential components that are in part independent (the philosophers are independent of each other and the chopsticks are independent of each other) and in part fully synchronized (every chopstick is fully synchronized with one of the two neighboring philosophers at a time). A deadlock state is known to arise if all the philosophers use the same protocol for getting the two chopsticks.

We suppose that every activity carried out by a philosopher has an exponentially distributed duration, with the exception of eating whose duration is assumed to follow a two-stage hypoexponential distribution with rates ε_1 and ε_2, respectively. The scenario can thus be described in MPC as follows:

$$((\mathit{Phil}_0/H_0) \parallel_\emptyset (\mathit{Phil}_1/H_1) \parallel_\emptyset \dots \parallel_\emptyset (\mathit{Phil}_{n-1}/H_{n-1})) \parallel_S (\mathit{Chop}_0 \parallel_\emptyset \mathit{Chop}_1 \parallel_\emptyset \dots \parallel_\emptyset \mathit{Chop}_{n-1})$$

where:

- Process constants Phil_i and Chop_i, $0 \leq i \leq n-1$, are used together with their defining equations in place of process variables and *rec* binders for the sake of readability.
- The synchronization set is given by $S = \{\mathit{get}_i, \mathit{put}_i \mid 0 \leq i \leq n-1\}$ and the rate synchronization operator $\otimes$ of Table 1 is assumed to be multiplication.
- The i-th chopstick, $0 \leq i \leq n-1$, behaves as follows:

$$\mathit{Chop}_i \triangleq {<}\mathit{get}_i, 1{>}.{<}\mathit{put}_i, 1{>}.\mathit{Chop}_i$$

n	#states original model	#trans. original model	#states reduced model	#trans. reduced model
2	26	42	22	36
3	124	297	100	243
4	626	$2,004$	466	$1,512$
5	$3,124$	$12,495$	$2,164$	$8,775$
6	$15,626$	$75,006$	$10,054$	$48,924$
7	$78,124$	$437,493$	$46,708$	$265,167$
8	$390,626$	$2,500,008$	$216,994$	$1,407,888$
9	$1,953,124$	$14,062,491$	$1,008,100$	$7,358,283$

Table 3: Size of the state space of the dining philosophers model before and after the reduction

- The i-th philosopher, $1 \leq i \leq n-1$, behaves as follows:

$$\begin{aligned} \mathit{Phil}_i \;\stackrel{\Delta}{=}\; & <\mathit{think}_i, \theta_i>.<\mathit{get}_i, \gamma_i>.<\mathit{get}_{(i+1) \bmod n}, \gamma_i>. \\ & <\mathit{eat\_first\_stage}_i, \varepsilon_{i,1}>.<\mathit{eat\_second\_stage}_i, \varepsilon_{i,2}>. \\ & <\mathit{put}_i, \varpi_i>.<\mathit{put}_{(i+1) \bmod n}, \varpi_i>.\mathit{Phil}_i \end{aligned}$$

  while in Phil_0 the order of the two actions *get* (and also of the two actions *put*) is reversed so to break symmetry and hence avoid deadlock.

- Each hiding set is given by $H_i = \{\mathit{eat\_first\_stage}_i, \mathit{eat\_second\_stage}_i\}$, $0 \leq i \leq n-1$.

Every sequential element of the form Phil_i/H_i, $0 \leq i \leq n-1$, can be minimized with respect to $\simeq_{\rm MB}$ by reducing its sequence of two τ-transitions of rates $\varepsilon_{i,1}$ and $\varepsilon_{i,2}$ to a single τ-transition of rate $\frac{\varepsilon_{i,1} \cdot \varepsilon_{i,2}}{\varepsilon_{i,1} + \varepsilon_{i,2}}$. The resulting $\simeq_{\rm MB}$-minimized sequential elements can then replace the original ones inside the overall parallel composition to derive a $\simeq_{\rm MB,g}$-equivalent state space that is smaller than the original one. In Table 3, we show the size of the state space before and after the reduction. For the considered values of n, the gain in terms of state space reduction tends to get close to 50%.

6. Related Work

The various weak Markovian bisimulation equivalences proposed in this paper are deeply inspired by *weak (Markovian) isomorphism* [22]. Its novelty was the introduction for the first time of the idea of reducing a sequence of exponentially timed τ-actions to a single exponentially timed τ-action preserving the expected duration of the action sequence, possibly in the context of exponentially timed visible actions.

Weak isomorphism was shown to be a congruence for both sequential and concurrent processes, and to be exact at stationary state only for processes satisfying a synchronization-related constraint analogous to the one in Thm. 4.22. Our constraint is formulated in a slightly more precise way than the one in [22], because it surely accounts for reducible computations of length greater than one interleaved with reducible computations of length one and, most importantly, it makes it clear that the initial state of a locally reducible computation can be reached via a non-joining synchronization neither in one step, nor in several steps.

In our work, we have revisited the idea at the basis of weak isomorphism by presenting it in two parts. We have first considered a weak Markovian bisimilarity that abstracts only from sequences of τ-transitions traversing fully unstable states, for which it is easier to illustrate the additional identifications made with respect to strong Markovian bisimilarity. The most important property of this equivalence is that of inducing a pseudo-aggregation that is exact at stationary state for all the considered processes; on the other hand, it is not a congruence with respect to parallel composition. We have then generalized our equivalence to retrieve compositionality over concurrent processes, but we have lost stationary-state exactness for a subset of processes. As a consequence, our study has emphasized the existence of a tradeoff between compositionality and exactness when abstracting from exponentially timed τ-transitions, and has caused a pseudo-aggregation unconditionally exact at stationary state like W-lumpability to emerge. Both facts went unnoticed in [22].

Moreover, we have extended the work of [22] in several directions:

- In lieu of isomorphism, we have considered the less restrictive and more useful framework of bisimulation. For example, processes such as $<a, \lambda_1>.P + <a, \lambda_2>.P$ and $<a, \lambda_1 + \lambda_2>.P$ cannot be identified under isomorphism thereby losing ordinary lumpability, which is instead included in W-lumpability and GW-lumpability.
- In place of individual sequences of exponentially timed τ-transitions, we have addressed trees of exponentially timed τ-transitions, and we have established the conditions under which such trees can be reduced by preserving both the expected duration and the execution probability of their branches. For instance, the pairs of process terms in Exs. 3.4 and 3.5 cannot be related by weak isomorphism.
- We have developed a sound and complete axiomatization that elucidates the fundamental equational laws at the basis of the proposed equivalences, as well as of weak isomorphism when restricting attention to sequences (rather than trees) of exponentially timed τ-transitions. The axiomatization has played a fundamental role in providing a structural definition of the induced pseudo-aggregations.
- We have addressed decidability issues. In particular, we have seen that our basic equivalence can be decided in polynomial time as long as there are no cycles of exponentially timed τ-transitions. Moreover, we have shown how to exploit the decision procedure for compositionally reducing the state space underlying concurrent processes according to the generalized equivalence.

Another approach to weakening bisimilarity in an exponentially timed setting comes from [14], where a variant of Markovian bisimilarity was defined that checks for exit rate equality with respect to all equivalence classes *apart from the one including the processes under examination*. This permits abstracting from exponentially timed selfloops, a property possessed also by the coarser variants of our equivalences, respectively introduced in Sects. 3.5 and 4.5 (see the rightmost part of Fig. 3). Congruence and axiomatization results were provided for the equivalence in [14], and a logical characterization based on CSL without the next operator was exhibited in [7]. Different from our study, exactness was not investigated.

We finally mention that our work shares several features with [1]. In that paper, a process calculus with durational actions was considered, where action durations are *fixed* instead of varying stochastically, such that the operational semantics takes into account both the starting time and the duration of each action execution. The authors then developed a *branching bisimilarity* capable of abstracting from the starting time and the duration of τ-transitions. As in our setting, it turns out that a single τ-action executed between two observable actions cannot be abstracted away, while a sequence of τ-transitions can be reduced to a single τ-transition whose duration is the sum of the durations of the original transitions. Notice that, in the deterministically timed setting of [1], the duration of each action corresponds to its expected value.

7. Conclusion

In this paper, we have presented weak variants of Markovian bisimilarity capable of reducing any sequence of at least two exponentially timed τ-transitions to a single exponentially timed τ-transition, whenever it is possible to preserve the expected duration and the execution probability of the sequence. Our study has revealed the existence of a tradeoff between compositionality and exactness. On the one hand, $\simeq_{\rm MB}$ is a congruence only over sequential processes, but induces a CTMC-level pseudo-aggregation that is exact at stationary state for all processes. On the other hand, $\simeq_{\rm MB,g}$ is a congruence over all processes, but induces a CTMC-level pseudo-aggregation that is exact at stationary state only for processes in which locally reducible computations are not directly or indirectly preceded by non-joining synchronizations.

From a behavioral equivalence viewpoint, this paper confirms, in a Markovian setting, the adequacy of the construction used in [27] for nondeterministic processes to single out the coarsest congruence contained in a weak bisimulation equivalence that is not compositional with respect to a choice operator. In a nondeterministic setting, different approaches to the definition of a weak bisimulation equivalence like branching bisimulation [18] and dynamic/progressing bisimulation [29] can be employed. However, they turn out to be no longer suitable in a Markovian setting, as they would be too demanding about matching exponentially timed τ-transitions.

On the stochastic side, we have assumed in this paper that an external observer can see the names of the actions that are performed by the processes, as well as the expected durations of those actions. Consequently, the external observer is not able to distinguish between an arbitrarily long sequence of exponentially timed τ-actions and a single exponentially timed τ-action having the same expected duration. This leads to a state space reduction that preserves reward-based performance measures at stationary state, but not reward-based performance measures at transient state, with the notable exception of those measures expressed in terms of the mean time to certain events.

We point out that considering higher moments of the duration of the actions – in addition to its expectation – may bring some advantage in terms of transient measure preservation. However, we would end up with a much finer weak Markovian bisimilarity, because the two random variables respectively quantifying the duration of a sequence of exponentially timed τ-transitions and the duration of a single exponentially timed τ-transition do not necessarily have the same values for higher moments when their expected values coincide. In particular, if we additionally take into account the variance, reductions of sequences of exponentially timed τ-transitions would no longer be admitted, as it would only be possible to change the order of the transitions in the sequence. For instance, if we consider the three process terms in Ex. 3.3, then $<\tau, \mu>.<\tau, \gamma>.Q$ and $<\tau, \gamma>.<\tau, \mu>.Q$ would still be identified, but neither could be reduced to $<\tau, \frac{\mu \cdot \gamma}{\mu + \gamma}>.Q$ anymore.

As far as future work is concerned, we would like to develop a sound and complete axiomatization also for $\simeq_{\mathrm{MB,g}}$. We expect this not to be a trivial task, as $\simeq_{\mathrm{MB,g}}$ looks at local computations that are interleaved with others, and we are not aware of the existence of axiomatizations for behavioral equivalences of this kind. Furthermore, we would like to investigate logical characterizations for the various weak variants of Markovian bisimilarity that we have proposed.

Finally, we plan to extend our approach to interactive Markov chains [20] and Markov automata [17], so to provide a means for merging sequences of exponentially distributed delays in an orthogonal-time setting. As in the integrated-time setting, for a process like $a \,.\, (\gamma) \,.\, b \,.\, \underline{0}$ the delay γ cannot be abstracted away, whereas $a \,.\, (\gamma_1) \,.\, \tau \,.\, (\gamma_2) \,.\, b \,.\, \underline{0}$ can be considered weakly equivalent to $a \,.\, (\frac{\gamma_1 \cdot \gamma_2}{\gamma_1 + \gamma_2}) \,.\, b \,.\, \underline{0}$. This may lead to find out a uniform definition of weak bisimilarity for exponentially timed processes and deterministically timed processes, as weak bisimilarity for the latter processes [36, 28] is precisely based on the idea of abstracting from τ-actions while summing up the intervening deterministic delays.

Acknowledgment: We are very grateful to the anonymous reviewers for their comments and suggestions. This work has been funded by MIUR-PRIN project *PaCo – Performability-Aware Computing: Logics, Models, and Languages.*

Appendix A. Proofs of Results

Proof of Prop. 3.7 (p. 8). Let us call P_1 and P_2 the two considered process terms and let $\sum_{i \in I} \mu_i \equiv \mu$ and $\sum_{j \in J_{i_1}} \gamma_{i_1,j} = \sum_{j \in J_{i_2}} \gamma_{i_2,j} \equiv \gamma$ for all $i_1, i_2 \in I$. It holds that $\mathcal{B} = \{(P_1, P_2), (P_2, P_1)\} \cup \{(P, P) \mid P \in \mathbb{P}\}$ is a weak Markovian bisimulation. In fact, for all $D \in \mathbb{P}_{\mathrm{nfu}}/\mathcal{B}$ there are three nontrivial cases all regarding P_1 and P_2:

- If D does not contain any $P_{i,j}$ and is not reachable via reducible computations from any $P_{i,j}$, then:
$$pbtm(P_1, D) \; = \; \emptyset \; = \; pbtm(P_2, D)$$

- If $D = \{P_{i_0,j_0}\}$ for some $i_0 \in I$ and $j_0 \in J_{i_0}$, then:
$$\begin{aligned} pbtm(P_1, D) \; &= \; \{\!| \, \frac{\mu_{i_0}}{\mu} \cdot \frac{\gamma_{i_0,j_0}}{\gamma} \cdot \left(\frac{1}{\mu} + \frac{1}{\gamma} \right) \, |\!\} \; = \\ &= \; \{\!| \, \frac{\frac{\mu_{i_0}}{\mu} \cdot \frac{\gamma_{i_0,j_0}}{\gamma} \cdot \left(\frac{1}{\mu} + \frac{1}{\gamma} \right)^{-1}}{r} \cdot \frac{1}{r} \, |\!\} \; = \; pbtm(P_2, D) \end{aligned}$$
where:
$$r \; = \; \sum_{i \in I} \sum_{j \in J_i} \frac{\mu_i}{\mu} \cdot \frac{\gamma_{i,j}}{\gamma} \cdot \left(\frac{1}{\mu} + \frac{1}{\gamma} \right)^{-1} \; = \; \left(\frac{1}{\mu} + \frac{1}{\gamma} \right)^{-1} \cdot \sum_{i \in I} \frac{\mu_i}{\mu} \cdot \sum_{j \in J_i} \frac{\gamma_{i,j}}{\gamma} \; = \; \left(\frac{1}{\mu} + \frac{1}{\gamma} \right)^{-1}$$

- If D is reachable via reducible computations from P_{i_0,j_0} for some $i_0 \in I$ and $j_0 \in J_{i_0}$, then for each such reducible computation the *probtime* contribution from P_1 to P_{i_0,j_0} coincides with the *probtime* contribution from P_2 to P_{i_0,j_0}:
$$\frac{\mu_{i_0}}{\mu} \cdot \frac{\gamma_{i_0,j_0}}{\gamma} \cdot \left(\frac{1}{\mu} + \frac{1}{\gamma}\right) \; = \; \frac{\frac{\mu_{i_0}}{\mu} \cdot \frac{\gamma_{i_0,j_0}}{\gamma} \cdot \left(\frac{1}{\mu} + \frac{1}{\gamma}\right)^{-1}}{r} \cdot \frac{1}{r}$$
where r is as in the previous case. Therefore:
$$pbtm(P_1, D) \; = \; pbtm(P_2, D) \qquad \blacksquare$$

Proof of Prop. 3.8 (p. 8). Let $P_1, P_2 \in \mathbb{P}$ be such that $P_1 \approx_{\rm MB} P_2$ and let $\mathcal{B}$ be a weak Markovian bisimulation containing the pair (P_1, P_2):

1. Let $P_1, P_2 \in \mathbb{P}_{\rm seq}$. Given $<a, \lambda> \in Act$, it turns out that the symmetric and transitive closure $\mathcal{B}'$ of the relation $\mathcal{B} \cup \{(<a, \lambda>.P_1, <a, \lambda>.P_2)\}$ is a weak Markovian bisimulation. In fact, there are two nontrivial cases regarding $<a, \lambda>.P_1$ and $<a, \lambda>.P_2$ and the equivalence class D with respect to $\mathcal{B}'$ such that $\{P_1, P_2\} \subseteq D$:
   - If $a \neq \tau$, then $<a, \lambda>.P_1, <a, \lambda>.P_2 \in \mathbb{P}_{\rm nfu}$ and for all $a' \in Name$ and $D' \in \mathbb{P}/\mathcal{B}'$ we have that:
   $$rate(<a, \lambda>.P_1, a', D') \; = \; rate(<a, \lambda>.P_2, a', D') \; = \; \left\{ \begin{array}{ll} \lambda & \text{if } a' = a \wedge D' = D \\ 0 & \text{if } a' \neq a \vee D' \neq D \end{array} \right.$$
   - If $a = \tau$, then $<a, \lambda>.P_1, <a, \lambda>.P_2 \in \mathbb{P}_{\rm fu}$ with the only equivalence class reachable in one step by both of them being D. Let $D' \in \mathbb{P}_{\rm nfu}/\mathcal{B}'$. If D' is not reachable from D via reducible computations, then:
   $$pbtm(<a, \lambda>.P_1, D') \; = \; \emptyset \; = \; pbtm(<a, \lambda>.P_2, D')$$
   otherwise for each reducible computation to D' the *probtime* contribution from $<a, \lambda>.P_1$ to D coincides with the *probtime* contribution from $<a, \lambda>.P_2$ to D (because $P_1 \approx_{\rm MB} P_2$) and hence:
   $$pbtm(<a, \lambda>.P_1, D') \; = \; pbtm(<a, \lambda>.P_2, D')$$
2. Given $H \subseteq Name_{\rm v}$, it turns out that the transitive closure $\mathcal{B}'$ of the relation $\mathcal{B} \cup \{(P'_1/H, P'_2/H) \mid (P'_1, P'_2) \in \mathcal{B}\}$ is a weak Markovian bisimulation. In fact, there are two nontrivial cases all regarding pairs $(P'_1/H, P'_2/H) \in \mathcal{B}'$ and equivalence classes D of the form $[P'/H]_{\mathcal{B}'} = \{P''/H \in \mathbb{P} \mid P'' \in [P']_{\mathcal{B}}\}$:
   - If $P'_1/H, P'_2/H \in \mathbb{P}_{\rm nfu}$, then $rate(P'_1/H, a, D) = rate(P'_2/H, a, D)$ because for $i = 1, 2$ it holds that:
   $$rate(P'_i/H, a, D) \; = \; \left\{ \begin{array}{ll} 0 & \text{if } a \in H \\ rate(P'_i, a, [P']_{\mathcal{B}}) & \text{if } a \notin H \cup \{\tau\} \\ \sum\limits_{b \in H \cup \{\tau\}} rate(P'_i, b, [P']_{\mathcal{B}}) & \text{if } a = \tau \end{array} \right.$$
   - If $P'_1/H, P'_2/H \in \mathbb{P}_{\rm fu}$, for $D \subseteq \mathbb{P}_{\rm nfu}$ we have that each reducible computation from P'_1 (resp. P'_2) to $[P']_{\mathcal{B}}$ induces a reducible computation from P'_1/H (resp. P'_2/H) to D with the same *probtime* measure. In addition, there might be further reducible computations from P'_1/H (resp. P'_2/H) to D originated from the fact that $\_/H$ has made some intermediate states between P'_1 (resp. P'_2) and $[P']_{\mathcal{B}}$ fully unstable. Since $(P'_1, P'_2) \in \mathcal{B}$ and $\mathcal{B}$ is a weak Markovian bisimulation, those intermediate states have to be pairwise related by $\mathcal{B}$ and hence have to pass the exit rate equality check. This is enough to guarantee that the multiset of additional reducible computations from P'_1/H to D having a certain expected duration, and the multiset of additional reducible computations from P'_2/H to D having the same expected duration, have the same sum of *probtime* measures. Therefore:
   $$pbtm(P'_1/H, D) \; = \; pbtm(P'_2/H, D) \qquad \blacksquare$$

Proof of Prop. 3.10 (p. 9). Let $P_1, P_2 \in \mathbb{P}$. The proof is divided into five parts:

- Firstly, we prove that $P_1 \sim_{\rm MB} P_2$ implies $P_1 \approx_{\rm MB} P_2$. If $P_1 \sim_{\rm MB} P_2$, then there exists a Markovian bisimulation $\mathcal{B}$ containing the pair (P_1, P_2). It turns out that $\mathcal{B}$ is a weak Markovian bisimulation too. In fact, observing that $\mathcal{B}$ cannot contain any pair composed of a fully unstable process term and a non-fully-unstable process term, the following holds whenever $(P'_1, P'_2) \in \mathcal{B}$:

– If $P'_1, P'_2 \in \mathbb{P}_{\mathrm{nfu}}$, then for all $a \in \mathit{Name}$ and $D \in \mathbb{P}/\mathcal{B}$:

$$\mathit{rate}(P'_1, a, D) \;=\; \mathit{rate}(P'_2, a, D)$$

The reason is that $(P'_1, P'_2) \in \mathcal{B}$ and $\mathcal{B}$ is a Markovian bisimulation.

– If $P'_1, P'_2 \in \mathbb{P}_{\mathrm{fu}}$, then for all $D \subseteq \mathbb{P}_{\mathrm{nfu}}/\mathcal{B}$:

$$\mathit{pbtm}(P'_1, D) \;=\; \mathit{pbtm}(P'_2, D)$$

The reason is that, since $(P'_1, P'_2) \in \mathcal{B}$ and $\mathcal{B}$ is a Markovian bisimulation, for each maximal multiset of reducible computations from P'_1 (resp. P'_2) to D whose corresponding traversed states form pairs contained in $\mathcal{B}$, there exists a maximal multiset of reducible computations from P'_2 (resp. P'_1) to D whose corresponding traversed states form pairs contained in $\mathcal{B}$, such that all corresponding states traversed by the reducible computations in the two multisets form pairs contained in $\mathcal{B}$. Therefore, the two multisets contribute to *pbtm* with the same sum of *probtime* measures.

- Secondly, we demonstrate that $P_1 \sim_{\mathrm{MB}} P_2$ implies $P_1 \simeq_{\mathrm{MB}} P_2$. Since we have proved that $\sim_{\mathrm{MB}} \subseteq \approx_{\mathrm{MB}}$, the equivalence classes of $\approx_{\mathrm{MB}}$ are unions of equivalence classes of $\sim_{\mathrm{MB}}$. Thus, if $P_1 \sim_{\mathrm{MB}} P_2$ and we take $a \in \mathit{Name}$ and $D \in \mathbb{P}/\approx_{\mathrm{MB}}$ with $D = \bigcup_{i \in I} D_i$ and $D_i \in \mathbb{P}/\sim_{\mathrm{MB}}$ for all $i \in I$, we have:

$$\mathit{rate}(P_1, a, D) \;=\; \sum_{i \in I} \mathit{rate}(P_1, a, D_i) \;=\; \sum_{i \in I} \mathit{rate}(P_2, a, D_i) \;=\; \mathit{rate}(P_2, a, D)$$

which means that $P_1 \simeq_{\mathrm{MB}} P_2$.

- Thirdly, we show that $P_1 \simeq_{\mathrm{MB}} P_2$ implies $P_1 \approx_{\mathrm{MB}} P_2$. Whenever $P_1 \simeq_{\mathrm{MB}} P_2$, then $P_1 \approx_{\mathrm{MB}} P_2$ because:

– If $P_1, P_2 \in \mathbb{P}_{\mathrm{nfu}}$, then for all $a \in \mathit{Name}$ and $D \in \mathbb{P}/\approx_{\mathrm{MB}}$:

$$\mathit{rate}(P_1, a, D) \;=\; \mathit{rate}(P_2, a, D)$$

The reason is that $P_1 \simeq_{\mathrm{MB}} P_2$.

– If $P_1, P_2 \in \mathbb{P}_{\mathrm{fu}}$, then for all $D \subseteq \mathbb{P}_{\mathrm{nfu}}/\approx_{\mathrm{MB}}$:

$$\mathit{pbtm}(P_1, D) \;=\; \mathit{pbtm}(P_2, D)$$

The reason is that, since $P_1 \simeq_{\mathrm{MB}} P_2$, both P_1 and P_2 reach in one step the same equivalence classes at the same rates and hence the first step towards D contributes to *pbtm* in the same way for P_1 and P_2. At that point, among those equivalence classes reached in one step by P_1 and P_2, it is enough to consider both D itself (if reachable in one step) and the classes from which it is possible to arrive at D via reducible computations.

- Fourthly, we prove that the inclusions are strict. For example, we have:

$$\begin{aligned} <a, \lambda>.<\tau, \mu>.<\tau, \gamma>.\underline{0} \;&\not\sim_{\mathrm{MB}}\; <a, \lambda>.<\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0} \\ <a, \lambda>.<\tau, \mu>.<\tau, \gamma>.\underline{0} \;&\simeq_{\mathrm{MB}}\; <a, \lambda>.<\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0} \end{aligned}$$

and:

$$\begin{aligned} <\tau, \mu>.<\tau, \gamma>.\underline{0} \;&\not\simeq_{\mathrm{MB}}\; <\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0} \\ <\tau, \mu>.<\tau, \gamma>.\underline{0} \;&\approx_{\mathrm{MB}}\; <\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0} \end{aligned}$$

- Finally, the fact that $P_1 \simeq_{\mathrm{MB}} P_2$ iff $P_1 \approx_{\mathrm{MB}} P_2$ when $P_1, P_2 \in \mathbb{P}_{\mathrm{nfu}}$ stems immediately from the definitions of $\simeq_{\mathrm{MB}}$ and $\approx_{\mathrm{MB}}$. ■

Proof of Thm. 3.12 (p. 9). Let $P_1, P_2 \in \mathbb{P}$ be such that $P_1 \simeq_{\mathrm{MB}} P_2$:

1. Let $P_1, P_2 \in \mathbb{P}_{\mathrm{seq}}$. By virtue of Prop. 3.10, from $P_1 \simeq_{\mathrm{MB}} P_2$ it follows that $P_1 \approx_{\mathrm{MB}} P_2$ and hence P_1 and P_2 belong to the same equivalence class D with respect to $\approx_{\mathrm{MB}}$. Given $<a, \lambda> \in \mathit{Act}$, for all $a' \in \mathit{Name}$ and $D' \in \mathbb{P}/\approx_{\mathrm{MB}}$ we have that:

$$\mathit{rate}(<a, \lambda>.P_1, a', D') \;=\; \mathit{rate}(<a, \lambda>.P_2, a', D') \;=\; \begin{cases} \lambda & \text{if } a' = a \wedge D' = D \\ 0 & \text{if } a' \neq a \vee D' \neq D \end{cases}$$

Therefore $<a, \lambda>.P_1 \simeq_{\mathrm{MB}} <a, \lambda>.P_2$.

2. Let $P_1, P_2 \in \mathbb{P}_{\text{seq}}$. Given $P \in \mathbb{P}_{\text{seq}}$, for all $a \in \mathit{Name}$ and $D \in \mathbb{P}/\approx_{\text{MB}}$ we have that:
$$\begin{aligned} \mathit{rate}(P_1 + P, a, D) \;&=\; \mathit{rate}(P_1, a, D) + \mathit{rate}(P, a, D) \;= \\ &=\; \mathit{rate}(P_2, a, D) + \mathit{rate}(P, a, D) \;=\; \mathit{rate}(P_2 + P, a, D) \end{aligned}$$
because $P_1 \simeq_{\text{MB}} P_2$. Therefore $P_1 + P \simeq_{\text{MB}} P_2 + P$ and $P + P_1 \simeq_{\text{MB}} P + P_2$.
3. Given $H \subseteq \mathit{Name}_{\text{v}}$, for all $a \in \mathit{Name}$ and $D \in \mathbb{P}/\approx_{\text{MB}}$ there are two cases:

   - If D does not contain any term of the form P/H, then:
   $$\mathit{rate}(P_1/H, a, D) \;=\; 0 \;=\; \mathit{rate}(P_2/H, a, D)$$
   - If $D = [P/H]_{\approx_{\text{MB}}}$, then we can exploit the congruence property of $\approx_{\text{MB}}$ with respect to the hiding operator as established by Prop. 3.8 in order to express D as $\bigcup_{P' \in D'} [P']_{\approx_{\text{MB}}}/H$, where D' is a maximal set including P of process terms that are pairwise not related by $\approx_{\text{MB}}$, such that $P'/H \approx_{\text{MB}} P/H$ for all $P' \in D'$. As a consequence, we have $\mathit{rate}(P_1/H, a, D) = \mathit{rate}(P_2/H, a, D)$ because for $i = 1, 2$ it holds that:
   $$\mathit{rate}(P_i/H, a, D) \;=\; \begin{cases} 0 & \text{if } a \in H \\ \sum\limits_{P' \in D'} \mathit{rate}(P_i, a, [P']_{\approx_{\text{MB}}}) & \text{if } a \notin H \cup \{\tau\} \\ \sum\limits_{P' \in D'} \; \sum\limits_{b \in H \cup \{\tau\}} \mathit{rate}(P_i, b, [P']_{\approx_{\text{MB}}}) & \text{if } a = \tau \end{cases}$$
   and $P_1 \simeq_{\text{MB}} P_2$.

   Therefore $P_1/H \simeq_{\text{MB}} P_2/H$. ■

Proof of Thm. 3.13 (p. 9). Let $P_1, P_2 \in \mathbb{P}_{\text{seq}}$. The proof is divided into two parts:

- $\Rightarrow$ If $P_1 \simeq_{\text{MB}} P_2$, then by virtue of Thm. 3.12 it follows that $P_1 + P \simeq_{\text{MB}} P_2 + P$ for all $P \in \mathbb{P}_{\text{seq}}$. Due to Prop. 3.10, this implies that $P_1 + P \approx_{\text{MB}} P_2 + P$ for all $P \in \mathbb{P}_{\text{seq}}$.
- $\Leftarrow$ Suppose that $P_1 + P \approx_{\text{MB}} P_2 + P$ for all $P \in \mathbb{P}_{\text{seq}}$. Since it is possible to find $\bar{P} \in \mathbb{P}_{\text{seq}}$ such that neither $P_1 + \bar{P}$ nor $P_2 + \bar{P}$ is fully unstable, from $P_1 + \bar{P} \approx_{\text{MB}} P_2 + \bar{P}$ it follows that $P_1 + \bar{P} \simeq_{\text{MB}} P_2 + \bar{P}$ because $\simeq_{\text{MB}}$ and $\approx_{\text{MB}}$ coincide over $\mathbb{P}_{\text{nfu}}$ as established by Prop. 3.10. Since for all $a \in \mathit{Name}$ and $D \in \mathbb{P}/\approx_{\text{MB}}$ it then holds that:
$$\begin{aligned} \mathit{rate}(P_1, a, D) \;&=\; \mathit{rate}(P_1 + \bar{P}, a, D) - \mathit{rate}(\bar{P}, a, D) \;= \\ &=\; \mathit{rate}(P_2 + \bar{P}, a, D) - \mathit{rate}(\bar{P}, a, D) \;=\; \mathit{rate}(P_2, a, D) \end{aligned}$$
we have that $P_1 \simeq_{\text{MB}} P_2$. ■

Proof of Prop. 3.16 (p. 9). Let $\mathcal{B}$ be a weak Markovian bisimulation up to $\approx_{\text{MB}}$. We first show that $(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\text{MB}})^+$ is a weak Markovian bisimulation by proving by induction on $n \in \mathbb{N}_{\geq 1}$ that, whenever $(P_1, P_2) \in (\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\text{MB}})^n$, then one of the following holds:

- $P_1, P_2 \in \mathbb{P}_{\text{nfu}}$ and for all $a \in \mathit{Name}$ and $D \in \mathbb{P}/(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\text{MB}})^+$:
$$\mathit{rate}(P_1, a, D) \;=\; \mathit{rate}(P_2, a, D)$$
- $P_1, P_2 \in \mathbb{P}_{\text{fu}}$ and for all $D \in \mathbb{P}_{\text{nfu}}/(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\text{MB}})^+$:
$$\mathit{pbtm}(P_1, D) \;=\; \mathit{pbtm}(P_2, D)$$

Let $(P_1, P_2) \in (\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\text{MB}})^n$:

- If $n = 1$, then $(P_1, P_2) \in \mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\text{MB}}$. There are two cases:
  - If $(P_1, P_2) \in \mathcal{B} \cup \mathcal{B}^{-1}$, then the result immediately follows from the fact that $\mathcal{B}$ is a weak Markovian bisimulation up to $\approx_{\text{MB}}$.
  - If $(P_1, P_2) \in \approx_{\text{MB}}$, then the result stems from the fact that $\approx_{\text{MB}} \subseteq (\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\text{MB}})^+$ and hence each equivalence class of $(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\text{MB}})^+$ is the union of some equivalence classes of $\approx_{\text{MB}}$.

- Let $n > 1$ and suppose that the result holds for all $(Q_1, Q_2) \in (\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB})^{n-1}$. From $(P_1, P_2) \in (\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB})^n$, we derive that there exists $P \in \mathbb{P}$ such that $(P_1, P) \in (\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB})^{n-1}$ and $(P, P_2) \in \mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB}$. Then the result holds both for the pair (P_1, P) – by the induction hypothesis – and for the pair (P, P_2) – by reasoning like in the case $n = 1$. As a consequence, the three process terms P_1, P_2, and P all belong either to $\mathbb{P}_{\rm nfu}$ or to $\mathbb{P}_{\rm fu}$, and hence the result follows for the pair (P_1, P_2) by transitivity of *rate* equality or *pbtm* equality, respectively.

Since we have proved that $(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB})^+$ is a weak Markovian bisimulation, $(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB})^+ \subseteq \approx_{\rm MB}$. On the other hand, $\mathcal{B} \subseteq (\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB})^+$. Therefore, $\mathcal{B} \subseteq \approx_{\rm MB}$ by transitivity of set inclusion, i.e., $(P_1, P_2) \in \mathcal{B}$ implies $P_1 \approx_{\rm MB} P_2$ for all $P_1, P_2 \in \mathbb{P}$. We also note that $(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB})^+ = \approx_{\rm MB}$. ■

Proof of Thm. 3.17 (p. 9). Without loss of generality, we assume for simplicity that the two sequential process terms $P_1, P_2 \in \mathcal{PL}$ such that $P_1 \simeq_{\rm MB} P_2$ contain free occurrences of a single process variable $X \in \mathit{Var}$. Consider the binary relation:

$$\mathcal{B} \; = \; \{(P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}) \mid P \in \mathcal{PL} \text{ sequential, free occurrences of } X \text{ at most}\}$$

which is a subset of $(\mathcal{PL}_{\rm nfu} \times \mathcal{PL}_{\rm nfu}) \cup (\mathcal{PL}_{\rm fu} \times \mathcal{PL}_{\rm fu})$. In fact, e.g., the case $P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\} \in \mathcal{PL}_{\rm nfu}$ and $P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\} \in \mathcal{PL}_{\rm fu}$ is not possible because:

- If P is not a process variable, then the actions enabled by $P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ and the actions enabled by $P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}$ coincide with the actions enabled by P.

- If P is a process variable, which must be X, then $P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ is equal to $\mathrm{rec}\, X : P_1$ and $P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}$ is equal to $\mathrm{rec}\, X : P_2$. The two resulting process terms are isomorphic to $P_1\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ and $P_2\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}$, respectively, with $P_1\{\mathrm{rec}\, X : P_1 \hookrightarrow X\} \simeq_{\rm MB} P_2\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}$ because $P_1 \simeq_{\rm MB} P_2$.

Similar to [27], we show that $\mathcal{B}$ has a property stronger than being a weak Markovian bisimulation up to $\approx_{\rm MB}$: for each sequential $P \in \mathcal{PL}$ containing free occurrences of X at most, it holds that for all action names $a \in \mathit{Name}$ and equivalence classes $D \in \mathcal{PL}/(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB})^+$:

$$\mathit{rate}(P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, a, D) \; \leq \; \mathit{rate}(P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}, a, D)$$

(like in [15], $\geq$ can be established between the two *rate* values with a symmetric argument, from which it can be concluded that the two *rate* values coincide).

If $\mathit{rate}(P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, a, D) = 0$, then the property trivially holds, otherwise we proceed by induction on the maximum depth $d \in \mathbb{N}_{\geq 1}$ of the inferences of the transitions from $P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ to D labeled with a:

- If $d = 1$, then only the semantic rule for the action prefix operator has been applied and hence P must be of the form $<a, \lambda>.P'$ (notice that it cannot be P equal to X because in that case $P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ would be equal to $\mathrm{rec}\, X : P_1$, which would contradict $d = 1$). Thus, for $i = 1, 2$ we have that $P\{\mathrm{rec}\, X : P_i \hookrightarrow X\}$ is of the form $<a, \lambda>.(P'\{\mathrm{rec}\, X : P_i \hookrightarrow X\})$. Since P' contains free occurrences of X at most, $(P'\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, P'\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}) \in \mathcal{B}$ and hence both process terms belong to D. Thus $\mathit{rate}(P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, a, D) = \lambda = \mathit{rate}(P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}, a, D)$.

- Let $d > 1$ and suppose that the property holds for all triples composed of a pair of process terms in $\mathcal{B}$, an equivalence class D', and an action name a', such that there are transitions from the first process term of the pair to D' labeled with a', and the maximum depth of their inferences is at most $d - 1$. We have the following cases:

  - If P is of the form $P' + P''$, then for $i = 1, 2$ we have that $P\{\mathrm{rec}\, X : P_i \hookrightarrow X\}$ is of the form $P'\{\mathrm{rec}\, X : P_i \hookrightarrow X\} + P''\{\mathrm{rec}\, X : P_i \hookrightarrow X\}$ and hence $\mathit{rate}(P\{\mathrm{rec}\, X : P_i \hookrightarrow X\}, a, D) = \mathit{rate}(P'\{\mathrm{rec}\, X : P_i \hookrightarrow X\}, a, D) + \mathit{rate}(P''\{\mathrm{rec}\, X : P_i \hookrightarrow X\}, a, D)$. In this case, the semantic rules for the alternative composition operator are applied first and hence the transitions from $P'\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ and $P''\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ to D labeled with a are considered (their inferences have maximum depth $d - 1$). If there are no such transitions from $P'\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$, then

$rate(P'\{\text{rec}\, X : P_1 \hookrightarrow X\}, a, D) = 0$, otherwise – since P' contains free occurrences of X at most – from the induction hypothesis it follows that $rate(P'\{\text{rec}\, X : P_1 \hookrightarrow X\}, a, D) \leq rate(P'\{\text{rec}\, X : P_2 \hookrightarrow X\}, a, D)$. Using a similar argument, we have that $rate(P''\{\text{rec}\, X : P_1 \hookrightarrow X\}, a, D) = 0$ or by the induction hypothesis $rate(P''\{\text{rec}\, X : P_1 \hookrightarrow X\}, a, D) \leq rate(P''\{\text{rec}\, X : P_2 \hookrightarrow X\}, a, D)$. Thus $rate(P\{\text{rec}\, X : P_1 \hookrightarrow X\}, a, D) \leq rate(P\{\text{rec}\, X : P_2 \hookrightarrow X\}, a, D)$.

– If P is a process variable, which must be X, then for $i = 1, 2$ we have that $P\{\text{rec}\, X : P_i \hookrightarrow X\}$ is equal to $\text{rec}\, X : P_i$, which in turn is isomorphic to $P_i\{\text{rec}\, X : P_i \hookrightarrow X\}$ and hence $rate(\text{rec}\, X : P_i, a, D) = rate(P_i\{\text{rec}\, X : P_i \hookrightarrow X\}, a, D)$. In this case, the semantic rule for recursion is applied first and hence the transitions from $P_1\{\text{rec}\, X : P_1 \hookrightarrow X\}$ to D labeled with a are considered (their inferences have maximum depth $d - 1$). Since P_1 contains free occurrences of X at most, from the induction hypothesis it follows that $rate(P_1\{\text{rec}\, X : P_1 \hookrightarrow X\}, a, D) \leq rate(P_1\{\text{rec}\, X : P_2 \hookrightarrow X\}, a, D)$, with $rate(P_1\{\text{rec}\, X : P_2 \hookrightarrow X\}, a, D) = rate(P_2\{\text{rec}\, X : P_2 \hookrightarrow X\}, a, D)$ because $P_1 \simeq_{\text{MB}} P_2$. Thus $rate(\text{rec}\, X : P_1, a, D) \leq rate(\text{rec}\, X : P_2, a, D)$.

– If P is of the form $\text{rec}\, Y : P'$, then there are two subcases:

  * If $Y = X$, then P contains no free occurrences of X. Therefore, for $i = 1, 2$ we have that $P\{\text{rec}\, X : P_i \hookrightarrow X\}$ is equal to P and hence $rate(P\{\text{rec}\, X : P_1 \hookrightarrow X\}, a, D) = rate(P\{\text{rec}\, X : P_2 \hookrightarrow X\}, a, D)$.
  * If $Y \neq X$, then for $i = 1, 2$ we have that $P\{\text{rec}\, X : P_i \hookrightarrow X\}$ is isomorphic to $P'\{\text{rec}\, Y : P' \hookrightarrow Y\}\{\text{rec}\, X : P_i \hookrightarrow X\}$ and hence $rate(P\{\text{rec}\, X : P_i \hookrightarrow X\}, a, D) = rate(P'\{\text{rec}\, Y : P' \hookrightarrow Y\}\{\text{rec}\, X : P_i \hookrightarrow X\}, a, D)$. In this case, the semantic rule for recursion is applied first and hence the transitions from $P'\{\text{rec}\, Y : P' \hookrightarrow Y\}\{\text{rec}\, X : P_1 \hookrightarrow X\}$ to D labeled with a are considered (their inferences have maximum depth $d - 1$). Since $P'\{\text{rec}\, Y : P' \hookrightarrow Y\}$ contains free occurrences of X at most, from the induction hypothesis it follows that $rate(P'\{\text{rec}\, Y : P' \hookrightarrow Y\}\{\text{rec}\, X : P_1 \hookrightarrow X\}, a, D) \leq rate(P'\{\text{rec}\, Y : P' \hookrightarrow Y\}\{\text{rec}\, X : P_2 \hookrightarrow X\}, a, D)$. Thus $rate(P\{\text{rec}\, X : P_1 \hookrightarrow X\}, a, D) \leq rate(P\{\text{rec}\, X : P_2 \hookrightarrow X\}, a, D)$.

From the property of $\mathcal{B}$ that we have proved (and the symmetrical property), it follows that $\mathcal{B}$ is a weak Markovian bisimulation up to $\approx_{\text{MB}}$. In fact, if $P\{\text{rec}\, X : P_1 \hookrightarrow X\}, P\{\text{rec}\, X : P_2 \hookrightarrow X\} \in \mathcal{PL}_{\text{fu}}$, then for all $D \in \mathcal{PL}_{\text{nfu}}/(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\text{MB}})^+$ it holds that $pbtm(P\{\text{rec}\, X : P_1 \hookrightarrow X\}, D) = pbtm(P\{\text{rec}\, X : P_2 \hookrightarrow X\}, D)$. The reason is that both $P\{\text{rec}\, X : P_1 \hookrightarrow X\}$ and $P\{\text{rec}\, X : P_2 \hookrightarrow X\}$ reach in one step the same equivalence classes at the same rates and hence the first step towards D contributes to $pbtm$ in the same way for $P\{\text{rec}\, X : P_1 \hookrightarrow X\}$ and $P\{\text{rec}\, X : P_2 \hookrightarrow X\}$.

Therefore, by virtue of Prop. 3.16 we have that $(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\text{MB}})^+ = \approx_{\text{MB}}$ and hence what we have proved is that, for each sequential $P \in \mathcal{PL}$ containing free occurrences of X at most, it holds that for all $a \in Name$ and $D \in \mathcal{PL}/\approx_{\text{MB}}$:

$$rate(P\{\text{rec}\, X : P_1 \hookrightarrow X\}, a, D) \;=\; rate(P\{\text{rec}\, X : P_2 \hookrightarrow X\}, a, D)$$

This means that $P\{\text{rec}\, X : P_1 \hookrightarrow X\} \simeq_{\text{MB}} P\{\text{rec}\, X : P_2 \hookrightarrow X\}$ for all sequential $P \in \mathcal{PL}$ containing free occurrences of X at most. We finally derive $\text{rec}\, X : P_1 \simeq_{\text{MB}} \text{rec}\, X : P_2$ by taking P equal to X. ■

Proof of Lemma 3.18 (p. 10). Let $P_1, P_2 \in \mathbb{P}_{\text{seq,nr}}$ be such that $P_1 \approx_{\text{MB}} P_2$ but $P_1 \not\simeq_{\text{MB}} P_2$. The proof is divided into three parts:

- Since $\approx_{\text{MB}}$ and $\simeq_{\text{MB}}$ coincide over $\mathbb{P}_{\text{nfu}}$ as established by Prop. 3.10, both P_1 and P_2 must be fully unstable. Since P_1 and P_2 are nonrecursive, no rec binder can occur in them and hence both of them must start with one or more alternative exponentially timed τ-actions, i.e., $P_1 \equiv \sum_{i \in I_1} {<}\tau, \mu_{1,i}{>}.P_{1,i}$ and $P_2 \equiv \sum_{i \in I_2} {<}\tau, \mu_{2,i}{>}.P_{2,i}$ where $I_1 \neq \emptyset, I_2 \neq \emptyset$ are finite index sets.

- If all the derivative process terms $P_{1,i}$, $i \in I_1$, of P_1 and $P_{2,i}$, $i \in I_2$, of P_2 were not fully unstable, then for all $k \in \{1, 2\}$ and $D \in \mathbb{P}_{\text{nfu}}/\approx_{\text{MB}}$ we would have:

  $$pbtm(P_k, D) \;=\; \{\!\!\{\, \frac{rate(P_k, \tau, D)}{rate_{\text{t}}(P_k)} \cdot \frac{1}{rate_{\text{t}}(P_k)} \,\}\!\!\}$$

  If we let:

$$v(P_k, D) \; = \; \frac{rate(P_k, \tau, D)}{rate_{\rm t}(P_k)} \cdot \frac{1}{rate_{\rm t}(P_k)}$$

we would derive:

$$rate(P_k, \tau, D) \; = \; v(P_k, D) \cdot rate_{\rm t}(P_k) \cdot rate_{\rm t}(P_k)$$

and also:

$$\sum_{D' \in \mathbb{P}_{\rm nfu}/\approx_{\rm MB}} v(P_k, D') \; = \; \frac{1}{rate_{\rm t}(P_k)} \cdot \sum_{D' \in \mathbb{P}_{\rm nfu}/\approx_{\rm MB}} \frac{rate(P_k, \tau, D')}{rate_{\rm t}(P_k)} \; = \; \frac{1}{rate_{\rm t}(P_k)}$$

or equivalently:

$$rate_{\rm t}(P_k) \; = \; 1/\sum_{D' \in \mathbb{P}_{\rm nfu}/\approx_{\rm MB}} v(P_k, D')$$

so that:

$$rate(P_k, \tau, D) \; = \; v(P_k, D)/\left(\sum_{D' \in \mathbb{P}_{\rm nfu}/\approx_{\rm MB}} v(P_k, D')\right)^2$$

From $P_1 \approx_{\rm MB} P_2$ and the fact that both P_1 and P_2 are fully unstable, for all $D \in \mathbb{P}_{\rm nfu}/\approx_{\rm MB}$ it would then follow that:

$$\begin{aligned} rate(P_1, \tau, D) \; &= \; v(P_1, D)/\left(\sum_{D' \in \mathbb{P}_{\rm nfu}/\approx_{\rm MB}} v(P_1, D')\right)^2 \; = \\ &= \; v(P_2, D)/\left(\sum_{D' \in \mathbb{P}_{\rm nfu}/\approx_{\rm MB}} v(P_2, D')\right)^2 \; = \; rate(P_2, \tau, D) \end{aligned}$$

while for $a \neq \tau$ or $D'' \in \mathbb{P}_{\rm fu}/\approx_{\rm MB}$ we would have:

$$rate(P_1, a, D'') \; = \; 0 \; = \; rate(P_2, a, D'')$$

In conclusion, $P_1 \not\approx_{\rm MB} P_2$ would be violated. Therefore, at least one process term belonging to $\{P_{1,i} \mid i \in I_1\} \cup \{P_{2,i} \mid i \in I_2\}$ must be fully unstable.

- If the two sets of equivalence classes with respect to $\approx_{\rm MB}$ reachable in one step by P_1 and P_2 were the same, say $\{D_1, D_2, \ldots, D_n\}$ with $n \in \mathbb{N}_{\geq 1}$, from $P_1 \approx_{\rm MB} P_2$ we would derive that for all $1 \leq i \leq n$:

$$\frac{rate(P_1, \tau, D_i)}{rate_{\rm t}(P_1)} \cdot \frac{1}{rate_{\rm t}(P_1)} \; = \; \frac{rate(P_2, \tau, D_i)}{rate_{\rm t}(P_2)} \cdot \frac{1}{rate_{\rm t}(P_2)}$$

and hence:

$$\sum_{i=1}^{n} \frac{rate(P_1, \tau, D_i)}{rate_{\rm t}(P_1)} \cdot \frac{1}{rate_{\rm t}(P_1)} \; = \; \sum_{i=1}^{n} \frac{rate(P_2, \tau, D_i)}{rate_{\rm t}(P_2)} \cdot \frac{1}{rate_{\rm t}(P_2)}$$

or equivalently:

$$\frac{1}{rate_{\rm t}(P_1)} \; = \; \frac{1}{rate_{\rm t}(P_2)}$$

As a consequence, for all $1 \leq i \leq n$ we would have:

$$rate(P_1, \tau, D_i) \; = \; rate(P_2, \tau, D_i)$$

while for $a \neq \tau$ or $D' \notin \{D_1, D_2, \ldots, D_n\}$ we would have:

$$rate(P_1, a, D') \; = \; 0 \; = \; rate(P_2, a, D')$$

In conclusion, $P_1 \not\approx_{\rm MB} P_2$ would be violated. Therefore, it must be $\{D \in \mathbb{P}/\approx_{\rm MB} \mid \exists i \in I_1 . \, P_{1,i} \in D\} \neq \{D \in \mathbb{P}/\approx_{\rm MB} \mid \exists i \in I_2 . \, P_{2,i} \in D\}$. ■

Proof of Prop. 3.19 (p. 11). Let $P_1, P_2 \in \mathbb{P}_{\rm seq,nr}$ be such that $P_1 \approx_{\rm MB} P_2$ but $P_1 \not\approx_{\rm MB} P_2$. By virtue of Lemma 3.18, it turns out that P_1 is of the form $\sum_{i \in I_1} {<}\tau, \mu_{1,i}{>}.P_{1,i}$ and P_2 is of the form $\sum_{i \in I_2} {<}\tau, \mu_{2,i}{>}.P_{2,i}$ where $I_1 \neq \emptyset, I_2 \neq \emptyset$ are finite index sets, and at least one process term in $\{P_{1,i} \mid i \in I_1\} \cup \{P_{2,i} \mid i \in I_2\}$ is fully unstable. The proof is divided into two parts:

- Firstly, we show that at least one of P_1 and P_2 is of the form $\sum_{i \in I} {<}\tau, \mu_i{>}. \sum_{j \in J_i} {<}\tau, \gamma_{i,j}{>}.P_{i,j}$ where $I \neq \emptyset$ is a finite index set and $J_i \neq \emptyset$ is a finite index set for all $i \in I$. There are two cases:
  - Suppose that at least one of P_1 and P_2 can reach in one step only a single equivalence class $D \in \mathbb{P}/\approx_{\rm MB}$. Assuming that it is P_1, there are two subcases:
    - ∗ If $D \subseteq \mathbb{P}_{\rm fu}$, then we immediately derive that P_1 is of the considered form, i.e., $P_1 \equiv \sum_{i \in I_1} {<}\tau, \mu_{1,i}{>}. \sum_{j \in J_i} {<}\tau, \gamma_{1,i,j}{>}.P_{1,i,j}$ where $J_i \neq \emptyset$ is a finite index set for all $i \in I_1$.

* If $D \subseteq \mathbb{P}_{\rm nfu}$, then by virtue of Lemma 3.18 it cannot be the only equivalence class with respect to $\approx_{\rm MB}$ reachable in one step by P_2. Since $D \subseteq \mathbb{P}_{\rm nfu}$ and D is the only equivalence class with respect to $\approx_{\rm MB}$ reachable in one step by P_1, from $P_1 \approx_{\rm MB} P_2$ it follows that all the other classes with respect to $\approx_{\rm MB}$ reachable in one step by P_2 must be subsets of $\mathbb{P}_{\rm fu}$ and lead only to D. Furthermore, P_2 cannot reach D in one step, because otherwise $pbtm(P_2, D)$ would contain at least two values whereas $pbtm(P_1, D)$ contains only $\frac{1}{rate_{\rm t}(P_1)}$, thus violating $P_1 \approx_{\rm MB} P_2$. Hence $P_2 \equiv \sum_{i \in I_2} {<}\tau, \mu_{2,i}{>}. \sum_{j \in J_i} {<}\tau, \gamma_{2,i,j}{>}.P_{2,i,j}$ where $J_i \neq \emptyset$ is a finite index set for all $i \in I_2$.

– Suppose that both P_1 and P_2 can reach in one step several equivalence classes with respect to $\approx_{\rm MB}$. The two behavioral equivalences $\approx_{\rm MB}$ and $\simeq_{\rm MB}$ differ only for the treatment of fully unstable process terms. In fact, the former equivalence applies a *pbtm*-based equality check to reducible computations, whereas the latter equivalence applies a *rate*-based equality check to initial transitions. As a consequence, from $P_1 \approx_{\rm MB} P_2$ and $P_1 \not\simeq_{\rm MB} P_2$ it follows that some reducible computations of one of P_1 and P_2 must necessarily occur reduced in the other one, with the reductions taking place at the beginning of those computations and preserving their execution probability and their expected duration (reductions preserving those quantities can occur at any stage of the considered computations as $P_1 \approx_{\rm MB} P_2$, but only the absence of such reductions taking place right at the beginning of the computations violates $P_1 \not\simeq_{\rm MB} P_2$).
To be precise, in addition to initial reductions, there is another reason – initial permutations – for which some reducible computations of one of P_1 and P_2 may differ at the beginning in the other one. In fact, since *pbtm* abstracts not only from the length of reducible computations, but also from the order of the exponentially timed τ-transitions forming those computations, some reducible computations of P_1 (resp. P_2) may occur in P_2 (resp. P_1) with the actions labeling their first transitions exchanged with the actions labeling their second transitions. However, since both P_1 and P_2 can reach in one step several equivalence classes with respect to $\approx_{\rm MB}$, any such permutation would lead either to the same process term – if the rates of the involved actions are the same – or to a process term not $\approx_{\rm MB}$-equivalent to the original one because of the alteration of initial action execution probabilities or of the expected sojourn time of the original process term and its derivatives – if the rates of the involved actions are different.
In general, when a reduction takes place at the beginning of a fully unstable process term P, the reduction cannot be concerned with a single reducible computation, but must involve all the reducible computations of P. The reason is that the reduction must preserve the execution probability and the expected duration of all the computations of P. In fact, since it takes place at the beginning of P, the reduction produces another fully unstable process term P' whose initial actions are slower than the initial actions of P. Therefore, the expected sojourn time of P' is necessarily greater than the expected sojourn time of P:

$$\frac{1}{rate_{\rm t}(P')} > \frac{1}{rate_{\rm t}(P)}$$

Now, if P had a computation that cannot be reduced because it contains a single exponentially timed τ-transition – say of rate $\mu \in \mathbb{R}_{>0}$ – ending up in a non-fully-unstable state – say belonging to $D \in \mathbb{P}/\approx_{\rm MB}$ – then that computation would have an execution probability in P' greater than its execution probability in P:

$$\frac{\mu}{rate_{\rm t}(P')} > \frac{\mu}{rate_{\rm t}(P)}$$

In order to avoid this alteration of the execution probability of the considered computation, in P' we should change the rate of the corresponding initial action from μ to $\frac{\mu}{rate_{\rm t}(P)}$ multiplied by the reciprocal of the expected duration of the other initial actions of P', but then we would increase the expected duration of the considered computation with respect to P. Therefore, in any case $pbtm(P', D)$ and $pbtm(P, D)$ would be different and hence it would turn out that $P' \not\approx_{\rm MB} P$.
As a consequence of the fact that all reducible computations of one of P_1 and P_2 must necessarily occur reduced at the beginning in the other one, at least one of P_1 and P_2 must be of the form $\sum_{i \in I} {<}\tau, \mu_i{>}. \sum_{j \in J_i} {<}\tau, \gamma_{i,j}{>}.P_{i,j}$ where $I \neq \emptyset$ is a finite index set and $J_i \neq \emptyset$ is a finite index set for all $i \in I$.

- Secondly, we show that the term P_1 or P_2 that is of the form $\sum_{i \in I} {<}\tau, \mu_i{>}. \sum_{j \in J_i} {<}\tau, \gamma_{i,j}{>}.P_{i,j}$ satisfies one of the two properties mentioned at the end of the proposition statement. For simplicity, we assume that only one of P_1 and P_2 is of that form. There are two cases:

  - If that term can reach in one step only a single equivalence class with respect to $\approx_{\rm MB}$, then we immediately derive that $\sum_{j \in J_{i_1}} {<}\tau, \gamma_{i_1,j}{>}.P_{i_1,j} \approx_{\rm MB} \sum_{j \in J_{i_2}} {<}\tau, \gamma_{i_2,j}{>}.P_{i_2,j}$ for all $i_1, i_2 \in I$.
  - If that term can reach in one step several equivalence classes with respect to $\approx_{\rm MB}$, then the property satisfied in the previous case cannot hold. However, as shown in the first part of the proof (see the second subcase of the first case for P_2 reaching several classes, as well as the second case), all the reducible computations of that term must necessarily occur reduced at the beginning of the other term, because $P_1 \approx_{\rm MB} P_2$ but $P_1 \not\simeq_{\rm MB} P_2$. This is possible iff $\sum_{j \in J_{i_1}} \gamma_{i_1,j} = \sum_{j \in J_{i_2}} \gamma_{i_2,j}$ for all $i_1, i_2 \in I$. In fact:
    * If the property above is satisfied, then by virtue of Prop. 3.7 all the computations of that term can be reduced at the beginning.
    * Suppose that all the computations of that term can be reduced at the beginning, but the property above is not satisfied. Then there exist $i_1, i_2 \in I$ such that $\sum_{j \in J_{i_1}} \gamma_{i_1,j} \neq \sum_{j \in J_{i_2}} \gamma_{i_2,j}$ and hence the derivative of ${<}\tau, \mu_{i_1}{>}$ and the derivative of ${<}\tau, \mu_{i_2}{>}$ have expected sojourn times different from each other. If we let $\mu = \sum_{i \in I} \mu_i$, $\gamma_1 = \sum_{j \in J_{i_1}} \gamma_{i_1,j}$, and $\gamma_2 = \sum_{j \in J_{i_2}} \gamma_{i_2,j}$, then it holds that $\gamma_1 \neq \gamma_2$ and $\frac{1}{\mu} + \frac{1}{\gamma_1} \neq \frac{1}{\mu} + \frac{1}{\gamma_2}$.
      Since all computations must be reduced at the beginning in a way that preserves their execution probability and their expected duration, these pieces of information must necessarily be part of the rates of the new initial exponentially timed τ-actions resulting from the reduction. In particular, the reduction of ${<}\tau, \mu_{i_1}{>}$ with ${<}\tau, \gamma_{i_1,j}{>}$, $j \in J_{i_1}$, gives rise to an exponentially timed τ-action whose rate is $\frac{\mu_{i_1}}{\mu} \cdot \frac{\gamma_{i_1,j}}{\gamma_1} \cdot (\frac{1}{\mu} + \frac{1}{\gamma_1})^{-1}$, whereas the reduction of ${<}\tau, \mu_{i_2}{>}$ with ${<}\tau, \gamma_{i_2,j}{>}$, $j \in J_{i_2}$, gives rise to an exponentially timed τ-action whose rate is $\frac{\mu_{i_2}}{\mu} \cdot \frac{\gamma_{i_2,j}}{\gamma_2} \cdot (\frac{1}{\mu} + \frac{1}{\gamma_2})^{-1}$.
      However, the resulting process term is not $\approx_{\rm MB}$-equivalent to the original one, which contradicts $P_1 \approx_{\rm MB} P_2$. In fact, while in the original process term the *probtime* of reaching $\{P_{i_1,j} \mid j \in J_{i_1}\}$ is $\frac{\mu_{i_1}}{\mu} \cdot (\frac{1}{\mu} + \frac{1}{\gamma_1})^{-1}$, in the process term resulting from the reduction it is the previous value divided by the square of the sum of the rates of the new initial exponentially timed τ-actions (see Ex. 3.6). This sum is not equal to $(\frac{1}{\mu} + \frac{1}{\gamma_1})^{-1}$, because each of its summands is given by a fraction of μ multiplied by $(\frac{1}{\mu} + \frac{1}{\gamma})^{-1}$ for some $\gamma \in \mathbb{R}_{>0}$, with two of these γ values – γ_1 and γ_2 – being different from each other. ■

Proof of Lemma 3.21 (p. 11). We proceed by induction on the syntactical structure of $P \in \mathbb{P}_{\rm seq,nr}$:

- If $P \equiv \underline{0}$, then the result follows by taking $Q \equiv \underline{0}$ (which is in $\simeq_{\rm MB}$-normal-form) and using reflexivity.
- If $P \equiv {<}a, \lambda{>}.P'$, then by the induction hypothesis there exists $Q' \in \mathbb{P}_{\rm seq,nr}$ in $\simeq_{\rm MB}$-normal-form such that $\mathcal{A}_{\rm MB} \vdash P' = Q'$. From substitutivity with respect to action prefix, we obtain that $\mathcal{A}_{\rm MB} \vdash {<}a, \lambda{>}.P' = {<}a, \lambda{>}.Q'$. There are two cases:
  - If ${<}a, \lambda{>}.Q'$ is in $\simeq_{\rm MB}$-normal-form, then we are done.
  - If ${<}a, \lambda{>}.Q'$ is not in $\simeq_{\rm MB}$-normal-form, then the result follows after applying $\mathcal{A}_{\rm MB,5}$ by virtue of transitivity.
- If $P \equiv P_1 + P_2$, then by the induction hypothesis there exist $Q_1, Q_2 \in \mathbb{P}_{\rm seq,nr}$ in $\simeq_{\rm MB}$-normal-form such that $\mathcal{A}_{\rm MB} \vdash P_1 = Q_1$ and $\mathcal{A}_{\rm MB} \vdash P_2 = Q_2$. From substitutivity with respect to alternative composition, we obtain that $\mathcal{A}_{\rm MB} \vdash P_1 + P_2 = Q_1 + Q_2$. There are two cases:
  - If $Q_1 + Q_2$ is in $\simeq_{\rm MB}$-normal-form, then we are done.

– If $Q_1 + Q_2$ is not in $\simeq_{\rm MB}$-normal-form, then the result follows after as many applications of $\mathcal{A}_{\rm MB,3}$ and $\mathcal{A}_{\rm MB,4}$ as needed – possibly preceded by applications of $\mathcal{A}_{\rm MB,1}$ and $\mathcal{A}_{\rm MB,2}$ – by virtue of substitutivity with respect to alternative composition as well as transitivity. ■

Proof of Thm. 3.22 (p. 11). The proof is divided into two parts:

$\Rightarrow$ The soundness part of the result comes from the following remarks:

- Since $\simeq_{\rm MB}$ is an equivalence relation and a congruence with respect to action prefix and alternative composition by virtue of Thm. 3.12, in any deduction based on $\mathcal{A}_{\rm MB}$ it is correct to use reflexivity, symmetry, transitivity, and substitutivity with respect to action prefix and alternative composition.
- The validity of axioms $\mathcal{A}_{\rm MB,1}$ to $\mathcal{A}_{\rm MB,4}$ – which are sound for $\sim_{\rm MB}$ – is ensured by $\sim_{\rm MB} \subsetneq \simeq_{\rm MB}$ as established by Prop. 3.10.
- The validity of axiom $\mathcal{A}_{\rm MB,5}$ stems from Props. 3.7 and 3.11.

$\Leftarrow$ Given $P_1, P_2 \in \mathbb{P}_{\rm seq,nr}$ such that $P_1 \simeq_{\rm MB} P_2$, we prove that $\mathcal{A}_{\rm MB} \vdash P_1 = P_2$ by assuming, without loss of generality, that both P_1 and P_2 are in $\simeq_{\rm MB}$-normal-form. In fact, if this were not the case, by virtue of Lemma 3.21 we could derive $Q_1, Q_2 \in \mathbb{P}_{\rm seq,nr}$ in $\simeq_{\rm MB}$-normal-form such that $\mathcal{A}_{\rm MB} \vdash P_1 = Q_1$ and $\mathcal{A}_{\rm MB} \vdash P_2 = Q_2$ (hence $P_1 \simeq_{\rm MB} Q_1$ and $P_2 \simeq_{\rm MB} Q_2$ due to the soundness of the axioms with respect to $\simeq_{\rm MB}$), with $Q_1 \simeq_{\rm MB} Q_2$ (because it also holds that $P_1 \simeq_{\rm MB} P_2$ and $\simeq_{\rm MB}$ is a transitive relation). So, if we proved $\mathcal{A}_{\rm MB} \vdash Q_1 = Q_2$ from $Q_1 \simeq_{\rm MB} Q_2$, then $\mathcal{A}_{\rm MB} \vdash P_1 = P_2$ would follow by transitivity. Let us proceed by induction on the syntactical structure of $P_1 \in \mathbb{P}_{\rm seq,nr}$ in $\simeq_{\rm MB}$-normal-form:

- If $P_1 \equiv \underline{0}$, then from $P_1 \simeq_{\rm MB} P_2$ and P_2 in $\simeq_{\rm MB}$-normal-form it follows that $P_2 \equiv \underline{0}$ too, hence the result by reflexivity.
- If $P_1 \equiv \sum_{i \in I_1} <a_i, \lambda_i>.P_{1,i}$ with I_1 finite and nonempty, then from $P_1 \simeq_{\rm MB} P_2$ and P_2 in $\simeq_{\rm MB}$-normal-form it follows that $P_2 \equiv \sum_{j \in I_2} <b_j, \mu_j>.P_{2,j}$ with I_2 finite and nonempty.
Moreover, for all $i, i' \in I_1$ such that $i \neq i'$ (resp. $j, j' \in I_2$ such that $j \neq j'$) it must hold that $a_i \neq a_{i'}$ or $P_{1,i} \not\simeq_{\rm MB} P_{1,i'}$ (resp. $b_j \neq b_{j'}$ or $P_{2,j} \not\simeq_{\rm MB} P_{2,j'}$). In fact, if it were $a_h = a_{h'}$ and $P_{1,h} \simeq_{\rm MB} P_{1,h'}$ for some $h, h' \in I_1$ such that $h \neq h'$, then by the induction hypothesis we would have $\mathcal{A}_{\rm MB} \vdash P_{1,h} = P_{1,h'}$ and hence $\mathcal{A}_{\rm MB} \vdash <a_h, \lambda_h>.P_{1,h} + <a_{h'}, \lambda_{h'}>.P_{1,h'} = <a_h, \lambda_h>.P_{1,h} + <a_h, \lambda_{h'}>.P_{1,h}$ by substitutivity, which would contradict the initial minimality of P_1 with respect to $\mathcal{A}_{\rm MB,4}$.
In addition, for all $i, i' \in I_1$ such that $i \neq i'$ (resp. $j, j' \in I_2$ such that $j \neq j'$) it must hold that $a_i \neq a_{i'}$ or $P_{1,i} \not\approx_{\rm MB} P_{1,i'}$ (resp. $b_j \neq b_{j'}$ or $P_{2,j} \not\approx_{\rm MB} P_{2,j'}$). In fact, if it were $a_h = a_{h'}$ and $P_{1,h} \approx_{\rm MB} P_{1,h'}$ for some $h, h' \in I_1$ such that $h \neq h'$, then $P_{1,h} \not\simeq_{\rm MB} P_{1,h'}$ and $P_{1,h} \approx_{\rm MB} P_{1,h'}$ would contradict the initial minimality of P_1 summand $<a_h, \lambda_h>.P_{1,h}$ or $<a_{h'}, \lambda_{h'}>.P_{1,h'}$ with respect to $\mathcal{A}_{\rm MB,5}$ by virtue of Prop. 3.19.
As a consequence, since $P_1 \simeq_{\rm MB} P_2$ and hence for all $a \in \mathit{Name}$ and $D \in \mathbb{P}/\approx_{\rm MB}$ we have that $\mathit{rate}(P_1, a, D) = \mathit{rate}(P_2, a, D)$, a bijective correspondence can be established between the set of summands of P_1 and the set of summands of P_2. For each summand $<a_i, \lambda_i>.P_{1,i}$ there exists exactly one summand $<b_j, \mu_j>.P_{2,j}$ such that $a_i = b_j$, $\lambda_i = \mu_j$, and $P_{1,i} \approx_{\rm MB} P_{2,j}$ – and hence $P_{1,i} \simeq_{\rm MB} P_{2,j}$ otherwise $<a_i, \lambda_i>.P_{1,i}$ or $<b_j, \mu_j>.P_{2,j}$ would not be initially minimal with respect to $\mathcal{A}_{\rm MB,5}$ by virtue of Prop. 3.19 – and vice versa. For each pair of corresponding summands $<a_i, \lambda_i>.P_{1,i}$ and $<b_j, \mu_j>.P_{2,j}$, from $P_{1,i} \simeq_{\rm MB} P_{2,j}$ and the induction hypothesis it follows that $\mathcal{A}_{\rm MB} \vdash P_{1,i} = P_{2,j}$ and hence $\mathcal{A}_{\rm MB} \vdash <a_i, \lambda_i>.P_{1,i} = <b_j, \mu_j>.P_{2,j}$ by substitutivity with respect to action prefix ($a_i = b_j$ and $\lambda_i = \mu_j$). Due to the bijectivity of the correspondence, we have $\mathcal{A}_{\rm MB} \vdash \sum_{i \in I_1} <a_i, \lambda_i>.P_{1,i} = \sum_{j \in I_2} <b_j, \mu_j>.P_{2,j}$ by substitutivity with respect to alternative composition. ■

Proof of Thm. 3.24 (p. 13). Of the three conditions in Def. 3.23, we consider only the second one, because the first one is ordinary lumpability, which is known to be exact at stationary state, and the third one can be derived through a double application of the rewriting rule in Fig. 2.

In the following, we use subscript l (resp. r) to denote probabilities related to the original (resp. aggregated) CTMC on the left (resp. right) of the rewriting rule depicted in Fig. 2.

With regard to the two stationary state probability vectors $\boldsymbol{\pi}_{\mathrm{l}}$ and $\boldsymbol{\pi}_{\mathrm{r}}$, we have that $\boldsymbol{\pi}_{\mathrm{l}}$ satisfies the following linear system of global balance equations for the original CTMC on the left:

$$\begin{array}{rcll} \pi_{\mathrm{l}}[s_i] \cdot \gamma & = & \pi_{\mathrm{l}}[s] \cdot \mu_i & \text{for every state } s_i \ (i \in I) \\ \pi_{\mathrm{l}}[s_{i,j}] \cdot E(s_{i,j}) & = & \pi_{\mathrm{l}}[s_i] \cdot \gamma_{i,j} + \sum\limits_{s'' \in S \setminus \{s, s_i, z\}} \pi_{\mathrm{l}}[s''] \cdot R(s'', s_{i,j}) & \text{for every state } s_{i,j} \ (i \in I, \, j \in J_i) \\ \pi_{\mathrm{l}}[s'] \cdot E(s') & = & \sum\limits_{s'' \in S} \pi_{\mathrm{l}}[s''] \cdot R(s'', s') & \text{for any other state } s' \text{ including } s \end{array}$$

while for the aggregated CTMC on the right we have that $\boldsymbol{\pi}_{\mathrm{r}}$ satisfies the following linear system of global balance equations:

$$\begin{array}{rcll} \pi_{\mathrm{r}}[s_{i,j}] \cdot E(s_{i,j}) & = & \pi_{\mathrm{r}}[z] \cdot \frac{\mu_i}{\mu} \cdot \frac{\gamma_{i,j}}{\gamma} \cdot \left(\frac{1}{\mu} + \frac{1}{\gamma} \right)^{-1} + \sum\limits_{s'' \in S \setminus \{s, s_i, z\}} \pi_{\mathrm{r}}[s''] \cdot R(s'', s_{i,j}) & \text{for every state } s_{i,j} \ (i \in I, \, j \in J_i) \\ \pi_{\mathrm{r}}[s'] \cdot E(s') & = & \sum\limits_{s'' \in S} \pi_{\mathrm{r}}[s''] \cdot R(s'', s') & \text{for any other state } s' \text{ including } z \end{array}$$

with both the $\pi_{\mathrm{l}}[.]$'s and the $\pi_{\mathrm{r}}[.]$'s summing up to 1. Since $\frac{1}{\mu} + \frac{1}{\gamma} = \frac{\mu + \gamma}{\mu \cdot \gamma}$, the second linear system is equivalent to:

$$\begin{array}{rcll} \pi_{\mathrm{r}}[s_{i,j}] \cdot E(s_{i,j}) & = & \pi_{\mathrm{r}}[z] \cdot \frac{\mu_i \cdot \gamma_{i,j}}{\mu + \gamma} + \sum\limits_{s'' \in S \setminus \{s, s_i, z\}} \pi_{\mathrm{r}}[s''] \cdot R(s'', s_{i,j}) & \text{for every state } s_{i,j} \ (i \in I, \, j \in J_i) \\ \pi_{\mathrm{r}}[s'] \cdot E(s') & = & \sum\limits_{s'' \in S} \pi_{\mathrm{r}}[s''] \cdot R(s'', s') & \text{for any other state } s' \text{ including } z \end{array}$$

Thanks to a new variable y replacing the set of $1 + |I|$ variables $\{\pi_{\mathrm{l}}[s]\} \cup \{\pi_{\mathrm{l}}[s_i] \mid i \in I\}$, we show that the system of global balance equations for the original CTMC on the left can be transformed into a linear system having the same number of variables and equations as well as the same coefficient matrix as the linear system equivalent to the system of global balance equations for the aggregated CTMC on the right.

By summing up over all $i \in I$ the first group of equations in the linear system for the original CTMC, we derive the following equation:

$$\sum_{i \in I} \pi_{\mathrm{l}}[s_i] \cdot \gamma \; = \; \pi_{\mathrm{l}}[s] \cdot \mu$$

If we let:

$$y \; = \; \pi_{\mathrm{l}}[s] + \sum_{i \in I} \pi_{\mathrm{l}}[s_i]$$

or equivalently:

$$\sum_{i \in I} \pi_{\mathrm{l}}[s_i] \; = \; y - \pi_{\mathrm{l}}[s]$$

then the last derived equation can be rewritten as follows:

$$y \cdot \gamma - \pi_{\mathrm{l}}[s] \cdot \gamma \; = \; \pi_{\mathrm{l}}[s] \cdot \mu$$

and hence:

$$\pi_{\mathrm{l}}[s] \; = \; y \cdot \frac{\gamma}{\mu + \gamma}$$

Since from the first group of equations in the linear system for the original CTMC we derive that for all $i \in I$ it holds that:

$$\pi_{\mathrm{l}}[s_i] \; = \; \pi_{\mathrm{l}}[s] \cdot \frac{\mu_i}{\gamma} \; = \; y \cdot \frac{\mu_i}{\mu + \gamma}$$

the second group of equations in the linear system for the original CTMC can be rewritten as follows:

$$\pi_{\mathrm{l}}[s_{i,j}] \cdot E(s_{i,j}) \; = \; y \cdot \frac{\mu_i \cdot \gamma_{i,j}}{\mu + \gamma} + \sum_{s'' \in S \setminus \{s, s_i, z\}} \pi_{\mathrm{l}}[s''] \cdot R(s'', s_{i,j})$$

In conclusion, the introduction of variable y causes the system of global balance equations for the original CTMC to be equivalent to the following one:

$$\begin{array}{rcll} \pi_{\mathrm{l}}[s_{i,j}] \cdot E(s_{i,j}) & = & y \cdot \frac{\mu_i \cdot \gamma_{i,j}}{\mu + \gamma} + \sum\limits_{s'' \in S \setminus \{s, s_i, z\}} \pi_{\mathrm{l}}[s''] \cdot R(s'', s_{i,j}) & \text{for every state } s_{i,j} \ (i \in I, \, j \in J_i) \\ \pi_{\mathrm{l}}[s'] \cdot E(s') & = & \sum\limits_{s'' \in S} \pi_{\mathrm{l}}[s''] \cdot R(s'', s') & \text{for any other state } s' \text{ including the one for } y \end{array}$$

with all the occurring $\pi_{\mathrm{l}}[.]$'s plus y summing up to 1, which has the same form as the linear system equivalent to the system of global balance equations for the aggregated CTMC. As a consequence:

$$\begin{aligned} y &= \pi_{\mathrm{r}}[z] & \\ \pi_{\mathrm{l}}[s_{i,j}] &= \pi_{\mathrm{r}}[s_{i,j}] & \text{for every state } s_{i,j}\ (i \in I,\ j \in J_i) \\ \pi_{\mathrm{l}}[s'] &= \pi_{\mathrm{r}}[s'] & \text{for any other state } s' \end{aligned}$$

from which stationary-state exactness follows because:

$$\pi_{\mathrm{r}}[z] \;=\; y \;=\; \pi_{\mathrm{l}}[s] + \sum_{i \in I} \pi_{\mathrm{l}}[s_i]$$

■

Proof of Prop. 4.8 (p. 20). Let $\mathbb{P}^{\mathrm{ndiv}}$ be the set of non-divergent process terms of $\mathbb{P}$ and consider $P_1, P_2 \in \mathbb{P}^{\mathrm{ndiv}}$. Suppose that $P_1 \approx_{\mathrm{MB}} P_2$ due to some weak Markovian bisimulation $\mathcal{B}$ over $\mathbb{P}^{\mathrm{ndiv}}$. We now prove that $\mathcal{B}$ is also a generalized weak Markovian bisimulation over $\mathbb{P}^{\mathrm{ndiv}}$, from which it will follow that $P_1 \approx_{\mathrm{MB,g}} P_2$. Given $(P'_1, P'_2) \in \mathcal{B}$, there are two cases:

- If $P'_1, P'_2 \in \mathbb{P}^{\mathrm{ndiv}}_{\mathrm{nfu}}$, then for $\mathcal{K}_{\mathrm{lr}}(P_1) = \mathcal{K}_{\mathrm{lr}}(P_1) \uplus \emptyset$ and $\mathcal{K}_{\mathrm{lr}}(P_2) = \mathcal{K}_{\mathrm{lr}}(P_2) \uplus \emptyset$ we have that:
  - For all $a \in \mathit{Name}$ and $D \in \mathbb{P}^{\mathrm{ndiv}}/\mathcal{B}$:
    $\mathit{rate}_{\mathrm{g}}(P'_1, a, D, \mathcal{K}'_{\mathrm{nlr}}(P'_1)) \;=\; \mathit{rate}(P'_1, a, D) \;=\; \mathit{rate}(P'_2, a, D) \;=\; \mathit{rate}_{\mathrm{g}}(P'_2, a, D, \mathcal{K}'_{\mathrm{nlr}}(P'_2))$
    because $\mathcal{B}$ is a weak Markovian bisimulation.
  - $\mathcal{K}''_{\mathrm{lr}}(P_1) = \emptyset = \mathcal{K}''_{\mathrm{lr}}(P_2)$ hence there is no $\mathit{pbtm}_{\mathrm{g}}$-based check to perform.
- If $P'_1, P'_2 \in \mathbb{P}^{\mathrm{ndiv}}_{\mathrm{fu}}$, then $\mathcal{K}_{\mathrm{nlr}}(P_1) = \emptyset = \mathcal{K}_{\mathrm{nlr}}(P_2)$ and for $\mathcal{K}_{\mathrm{lr}}(P_1) = \emptyset \uplus \mathcal{K}_{\mathrm{lr}}(P_1)$ and $\mathcal{K}_{\mathrm{lr}}(P_2) = \emptyset \uplus \mathcal{K}_{\mathrm{lr}}(P_2)$ we have that:
  - For all $a \in \mathit{Name}$ and $D \in \mathbb{P}^{\mathrm{ndiv}}/\mathcal{B}$:
    $\mathit{rate}_{\mathrm{g}}(P'_1, a, D, \mathcal{K}'_{\mathrm{nlr}}(P'_1)) \;=\; 0 \;=\; \mathit{rate}_{\mathrm{g}}(P'_2, a, D, \mathcal{K}'_{\mathrm{nlr}}(P'_2))$
  - For $i = 1, 2$, it holds that P'_i has a single tree of locally reducible computations with respect to $\mathcal{K}_{\mathrm{lr}}(P'_i)$, with each such computation terminating with a non-fully-unstable state because P'_i is not divergent. Given $D \in \mathbb{P}^{\mathrm{ndiv}}/\mathcal{B}$, there are two subcases:
    - If $D \subseteq \mathbb{P}^{\mathrm{ndiv}}_{\mathrm{nfu}}$, then:
      $\mathit{pbtm}_{\mathrm{g}}(P'_1, D, \mathcal{K}_{\mathrm{lr}}(P'_1)) \;=\; \mathit{pbtm}(P'_1, D) \;=\; \mathit{pbtm}(P'_2, D) \;=\; \mathit{pbtm}_{\mathrm{g}}(P'_2, D, \mathcal{K}_{\mathrm{lr}}(P'_2))$
      because $\mathcal{B}$ is a weak Markovian bisimulation.
    - If $D \subseteq \mathbb{P}^{\mathrm{ndiv}}_{\mathrm{fu}}$, then:
      $\mathit{pbtm}_{\mathrm{g}}(P'_1, D, \mathcal{K}_{\mathrm{lr}}(P'_1)) \;=\; \emptyset \;=\; \mathit{pbtm}_{\mathrm{g}}(P'_2, D, \mathcal{K}_{\mathrm{lr}}(P'_2))$
      because D cannot contain any final state of a locally reducible computation. ■

Proof of Prop. 4.10 (p. 20). Let $P_1, P_2 \in \mathbb{P}$ be such that $P_1 \approx_{\mathrm{MB,g}} P_2$ and let $\mathcal{B}$ be a generalized weak Markovian bisimulation containing the pair (P_1, P_2):

1. Let $P_1, P_2 \in \mathbb{P}_{\mathrm{seq}}$. Given $<a, \lambda> \in \mathit{Act}$, it turns out that the relation $\mathcal{B}' = \mathcal{B} \cup \{(<a, \lambda>.P_1, <a, \lambda>.P_2), (<a, \lambda>.P_2, <a, \lambda>.P_1)\}$ is a generalized weak Markovian bisimulation. In fact, there are two nontrivial cases regarding $<a, \lambda>.P_1$ and $<a, \lambda>.P_2$ and the equivalence class D with respect to $\mathcal{B}'^{+}$ such that $\{P_1, P_2\} \subseteq D$:
   - If $a \neq \tau$, then $\mathcal{K}_{\mathrm{lr}}(<a, \lambda>.P_1) = \mathcal{K}_{\mathrm{lr}}(<a, \lambda>.P_2) = \emptyset$, hence $\mathcal{K}'_{\mathrm{nlr}}(<a, \lambda>.P_1) = \mathcal{K}_{\mathrm{nlr}}(<a, \lambda>.P_1)$ and $\mathcal{K}'_{\mathrm{nlr}}(<a, \lambda>.P_2) = \mathcal{K}_{\mathrm{nlr}}(<a, \lambda>.P_2)$. For all $a' \in \mathit{Name}$ and $D' \in \mathbb{P}/\mathcal{B}'^{+}$, we have that:
     $$\begin{aligned} \mathit{rate}_{\mathrm{g}}(<a, \lambda>.P_1, a', D', \mathcal{K}'_{\mathrm{nlr}}(<a, \lambda>.P_1)) &= \mathit{rate}_{\mathrm{g}}(<a, \lambda>.P_2, a', D', \mathcal{K}'_{\mathrm{nlr}}(<a, \lambda>.P_2)) \\ &= \begin{cases} \lambda & \text{if } a' = a \wedge D' = D \\ 0 & \text{if } a' \neq a \vee D' \neq D \end{cases} \end{aligned}$$
   - If $a = \tau$, then $\mathcal{K}_{\mathrm{nlr}}(<a, \lambda>.P_1) = \mathcal{K}_{\mathrm{nlr}}(<a, \lambda>.P_2) = \emptyset$. With the factorization $\mathcal{K}_{\mathrm{lr}}(<a, \lambda>.P_1) = \emptyset \uplus \mathcal{K}_{\mathrm{lr}}(<a, \lambda>.P_1)$ and $\mathcal{K}_{\mathrm{lr}}(<a, \lambda>.P_2) = \emptyset \uplus \mathcal{K}_{\mathrm{lr}}(<a, \lambda>.P_2)$, for all $a' \in \mathit{Name}$ and $D' \in \mathbb{P}/\mathcal{B}'^{+}$ we have that:
     $\mathit{rate}_{\mathrm{g}}(<a, \lambda>.P_1, a', D', \mathcal{K}'_{\mathrm{nlr}}(<a, \lambda>.P_1)) \;=\; 0 \;=\; \mathit{rate}_{\mathrm{g}}(<a, \lambda>.P_2, a', D', \mathcal{K}'_{\mathrm{nlr}}(<a, \lambda>.P_2))$
     Each of $<a, \lambda>.P_1$ and $<a, \lambda>.P_2$ has precisely one tree of locally reducible computation with respect to the entire set $\{1\}$ of its positions. There are two subcases:

– If $P_1, P_2 \in \mathbb{P}_{\rm nfu}$, then P_1 and P_2 are the final states of the two locally reducible computations, respectively, and for all $D' \in \mathbb{P}/\mathcal{B}'^+$ we have that:

$$pbtm_{\rm g}(<a,\lambda>.P_1, D', \{1\}) \; = \; pbtm_{\rm g}(<a,\lambda>.P_2, D', \{1\}) \; = \; \begin{cases} \{\!| \frac{1}{\lambda} |\!\} & \text{if } D' = D \\ \emptyset & \text{if } D' \neq D \end{cases}$$

– If $P_1, P_2 \in \mathbb{P}_{\rm fu}$, then from $(P_1, P_2) \in \mathcal{B}$ and the fact that $\mathcal{B}$ is a generalized weak Markovian bisimulation it follows that for each computation of P_1 locally reducible with respect to the entire position set $\{1\}$ there exists a computation of P_2 locally reducible with respect to the entire position set $\{1\}$ such that for all $D' \in \mathbb{P}/\mathcal{B}'^+$:

$$pbtm_{\rm g}(P_1, D', \{1\}) \; = \; pbtm_{\rm g}(P_2, D', \{1\})$$

and vice versa. Therefore, the two trees of locally reducible computations of $<a,\lambda>.P_1$ and $<a,\lambda>.P_2$ are identical backward extensions of the trees of locally reducible computations of P_1 and P_2, respectively. As a consequence, for all $D' \in \mathbb{P}/\mathcal{B}'^+$ we have that:

$$pbtm_{\rm g}(<a,\lambda>.P_1, D', \{1\}) \; = \; pbtm_{\rm g}(<a,\lambda>.P_2, D', \{1\})$$

because for $i = 1, 2$ the multiset $pbtm_{\rm g}(<a,\lambda>.P_i, D', \{1\})$ is obtained from the multiset $pbtm_{\rm g}(P_i, D', \{1\})$ by adding the expected duration $\frac{1}{\lambda}$ of the only exponentially timed τ-transition departing from $<a,\lambda>.P_i$ to the second factor of each $probtime_{\rm g}$ value contained in the latter multiset.

2. Given $H \subseteq Name_{\rm v}$, it turns out that the relation $\mathcal{B}' = \mathcal{B} \cup \{(P'_1/H, P'_2/H), (P'_2/H, P'_1/H) \mid (P'_1, P'_2) \in \mathcal{B}\}$ is a generalized weak Markovian bisimulation. The only nontrivial cases regard pairs $(P'_1/H, P'_2/H) \in \mathcal{B}'$ and equivalence classes D of the form $[P'/H]_{\mathcal{B}'^+} = \{P''/H \in \mathbb{P} \mid P'' \in [P']_{\mathcal{B}^+}\}$.
Let $\mathcal{K}_{\rm lr}(P'_1) = \mathcal{K}'_{\rm lr}(P'_1) \uplus \mathcal{K}''_{\rm lr}(P'_1)$ and $\mathcal{K}_{\rm lr}(P'_2) = \mathcal{K}'_{\rm lr}(P'_2) \uplus \mathcal{K}''_{\rm lr}(P'_2)$ be the factorization under which (P'_1, P'_2) belongs to the generalized weak Markovian bisimulation $\mathcal{B}$. Observing that $\mathcal{K}(P'_1/H) = \mathcal{K}(P'_1)$ and $\mathcal{K}(P'_2/H) = \mathcal{K}(P'_2)$, with the factorization $\mathcal{K}_{\rm lr}(P'_1/H) = \mathcal{K}'_{\rm lr}(P'_1) \uplus \mathcal{K}''_{\rm lr}(P'_1)$ and $\mathcal{K}_{\rm lr}(P'_2/H) = \mathcal{K}'_{\rm lr}(P'_2) \uplus \mathcal{K}''_{\rm lr}(P'_2)$ we have that, for all $a \in Name$, if $a \notin H$ then:

$$\begin{aligned} rate_{\rm g}(P'_1/H, a, D, \mathcal{K}'_{\rm nlr}(P'_1/H)) & = rate_{\rm g}(P'_1, a, [P']_{\mathcal{B}^+}, \mathcal{K}'_{\rm nlr}(P'_1)) \\ & = rate_{\rm g}(P'_2, a, [P']_{\mathcal{B}^+}, \mathcal{K}'_{\rm nlr}(P'_2)) \\ & = rate_{\rm g}(P'_2/H, a, D, \mathcal{K}'_{\rm nlr}(P'_2/H)) \end{aligned}$$

otherwise:

$$\begin{aligned} rate_{\rm g}(P'_1/H, \tau, D, \mathcal{K}'_{\rm nlr}(P'_1/H)) & = \sum_{b \in H \cup \{\tau\}} rate_{\rm g}(P'_1, b, [P']_{\mathcal{B}^+}, \mathcal{K}'_{\rm nlr}(P'_1)) \\ & = \sum_{b \in H \cup \{\tau\}} rate_{\rm g}(P'_2, b, [P']_{\mathcal{B}^+}, \mathcal{K}'_{\rm nlr}(P'_2)) \\ & = rate_{\rm g}(P'_2/H, \tau, D, \mathcal{K}'_{\rm nlr}(P'_2/H)) \end{aligned}$$

Moreover, since $(P'_1, P'_2) \in \mathcal{B}$ and $\mathcal{B}$ is a generalized weak Markovian bisimulation, for each computation of P'_1 locally reducible with respect to $K_1 \subseteq \mathcal{K}''_{\rm lr}(P'_1)$ there exists a computation of P'_2 locally reducible with respect to $K_2 \subseteq \mathcal{K}''_{\rm lr}(P'_2)$ such that:

$$pbtm_{\rm g}(P'_1, [P']_{\mathcal{B}^+}, K_1) \; = \; pbtm_{\rm g}(P'_2, [P']_{\mathcal{B}^+}, K_2)$$

and vice versa. For $i = 1, 2$, consider the computation of P'_i/H locally reducible with respect to K_i obtained from the computation of P'_i locally reducible with respect to K_i by applying "$\_/H$" to all the states traversed by the latter computation. The former computation thus coincides (up to state hiding) with the latter, or extends it with further exponentially timed τ-transitions such that their corresponding exponentially timed H-labeled transitions depart from and go to states related by the generalized weak Markovian bisimulation $\mathcal{B}$. Therefore:

$$pbtm_{\rm g}(P'_1/H, D, K_1) \; = \; pbtm_{\rm g}(P'_2/H, D, K_2)$$

3. Given $S \subseteq Name_{\rm v}$, it turns out that the relation $\mathcal{B}' = \mathcal{B} \cup \{(P'_1 \,\|_S\, P, P'_2 \,\|_S\, P), (P'_2 \,\|_S\, P, P'_1 \,\|_S\, P) \mid (P'_1, P'_2) \in \mathcal{B} \wedge P \in \mathbb{P}\}$ is a generalized weak Markovian bisimulation. The only nontrivial cases regard pairs $(P'_1 \,\|_S\, P, P'_2 \,\|_S\, P) \in \mathcal{B}'$ and equivalence classes D of the form $[P' \,\|_S\, Q]_{\mathcal{B}'^+} = \{P'' \,\|_S\, Q \in \mathbb{P} \mid P'' \in [P']_{\mathcal{B}^+}\}$.
Let $\mathcal{K}_{\rm lr}(P'_1) = \mathcal{K}'_{\rm lr}(P'_1) \uplus \mathcal{K}''_{\rm lr}(P'_1)$ and $\mathcal{K}_{\rm lr}(P'_2) = \mathcal{K}'_{\rm lr}(P'_2) \uplus \mathcal{K}''_{\rm lr}(P'_2)$ be the factorization under which (P'_1, P'_2) belongs to the generalized weak Markovian bisimulation $\mathcal{B}$. Observing that $\mathcal{K}(P'_1 \,\|_S\, P) = \mathcal{K}(P'_1) \uplus \mathcal{K}(P)$ and $\mathcal{K}(P'_2 \,\|_S\, P) = \mathcal{K}(P'_2) \uplus \mathcal{K}(P)$, with the factorization $\mathcal{K}_{\rm lr}(P'_1 \,\|_S\, P) = (\mathcal{K}'_{\rm lr}(P'_1) \uplus$

$\mathcal{K}'_{\mathrm{lr}}(P)) \uplus (\mathcal{K}''_{\mathrm{lr}}(P'_1) \uplus \mathcal{K}''_{\mathrm{lr}}(P))$ and $\mathcal{K}_{\mathrm{lr}}(P'_2 \,\|_S\, P) = (\mathcal{K}'_{\mathrm{lr}}(P'_2) \uplus \mathcal{K}'_{\mathrm{lr}}(P)) \uplus (\mathcal{K}''_{\mathrm{lr}}(P'_2) \uplus \mathcal{K}''_{\mathrm{lr}}(P))$ we have that for all $a \in \mathit{Name}$:

$$\mathit{rate}_{\mathrm{g}}(P'_1 \,\|_S\, P, a, D, \mathcal{K}'_{\mathrm{nlr}}(P'_1 \,\|_S\, P)) \; = \; \mathit{rate}_{\mathrm{g}}(P'_2 \,\|_S\, P, a, D, \mathcal{K}'_{\mathrm{nlr}}(P'_2 \,\|_S\, P))$$

because for $i = 1, 2$ it holds that $\mathit{rate}(P'_i \,\|_S\, P, a, D, \mathcal{K}'_{\mathrm{nlr}}(P'_i \,\|_S\, P))$ is equal to:

$$\begin{cases} \mathit{rate}_{\mathrm{g}}(P'_i, a, [P']_{\mathcal{B}+}, \mathcal{K}'_{\mathrm{nlr}}(P'_i)) \otimes \mathit{rate}_{\mathrm{g}}(P, a, \{Q\}, \mathcal{K}'_{\mathrm{nlr}}(P)) & \text{if } a \in S \\ \mathit{rate}_{\mathrm{g}}(P'_i, a, [P']_{\mathcal{B}+}, \mathcal{K}'_{\mathrm{nlr}}(P'_i)) + \mathit{rate}_{\mathrm{g}}(P, a, \{Q\}) & \text{if } a \notin S \wedge P'_i \in [P']_{\mathcal{B}+} \wedge P \equiv Q \\ \mathit{rate}_{\mathrm{g}}(P'_i, a, [P']_{\mathcal{B}+}, \mathcal{K}'_{\mathrm{nlr}}(P'_i)) & \text{if } a \notin S \wedge P'_i \notin [P']_{\mathcal{B}+} \wedge P \equiv Q \\ \mathit{rate}_{\mathrm{g}}(P, a, \{Q\}, \mathcal{K}'_{\mathrm{nlr}}(P)) & \text{if } a \notin S \wedge P'_i \in [P']_{\mathcal{B}+} \wedge P \not\equiv Q \\ 0 & \text{if } a \notin S \wedge P'_i \notin [P']_{\mathcal{B}+} \wedge P \not\equiv Q \end{cases}$$

Moreover, since $(P'_1, P'_2) \in \mathcal{B}$ and $\mathcal{B}$ is a generalized weak Markovian bisimulation, for each computation of P'_1 locally reducible with respect to $K_1 \subseteq \mathcal{K}''_{\mathrm{lr}}(P'_1)$ there exists a computation of P'_2 locally reducible with respect to $K_2 \subseteq \mathcal{K}''_{\mathrm{lr}}(P'_2)$ such that:

$$\mathit{pbtm}_{\mathrm{g}}(P'_1, [P']_{\mathcal{B}+}, K_1) \; = \; \mathit{pbtm}_{\mathrm{g}}(P'_2, [P']_{\mathcal{B}+}, K_2)$$

and vice versa. For $i = 1, 2$, consider the computation of $P'_i \,\|_S\, P$ locally reducible with respect to K_i obtained from the computation of P'_i locally reducible with respect to K_i by applying "$\_ \,\|_S\, P$" to all the states traversed by the latter computation. The former computation thus coincides (up to state parallel composition and the consequent possible variation of the context) with the latter, or extends it with further exponentially timed τ-transitions – emerged because alternative transitions labeled with exponentially timed visible actions have been cut off due to unsatisfied synchronization constraints – that depart from and go to states related by the generalized weak Markovian bisimulation $\mathcal{B}$. Therefore:

$$\mathit{pbtm}_{\mathrm{g}}(P'_1 \,\|_S\, P, D, K_1) \; = \; \mathit{pbtm}_{\mathrm{g}}(P'_2 \,\|_S\, P, D, K_2)$$

A similar reasoning applies if P has a computation locally reducible with respect to $K \subseteq \mathcal{K}''_{\mathrm{lr}}(P)$. ■

Proof of Prop. 4.12 (p. 21). Let $P_1, P_2 \in \mathbb{P}$. The proof is divided into five parts:

- Firstly, we prove that $P_1 \sim_{\mathrm{MB}} P_2$ implies $P_1 \approx_{\mathrm{MB,g}} P_2$. If $P_1 \sim_{\mathrm{MB}} P_2$, then there exists a Markovian bisimulation $\mathcal{B}$ containing the pair (P_1, P_2). It turns out that $\mathcal{B}$ is a generalized weak Markovian bisimulation too. In fact, the following holds whenever $(P'_1, P'_2) \in \mathcal{B}$ under the factorization $\mathcal{K}_{\mathrm{lr}}(P'_1) = \emptyset \uplus \mathcal{K}_{\mathrm{lr}}(P'_1)$ and $\mathcal{K}_{\mathrm{lr}}(P'_2) = \emptyset \uplus \mathcal{K}_{\mathrm{lr}}(P'_2)$:

  - For all $a \in \mathit{Name}$ and $D \in \mathbb{P}/\mathcal{B}$:
  $$\mathit{rate}_{\mathrm{g}}(P'_1, a, D, \mathcal{K}'_{\mathrm{nlr}}(P'_1)) \; = \; \mathit{rate}(P'_1, a, D) \; = \; \mathit{rate}(P'_2, a, D) \; = \; \mathit{rate}_{\mathrm{g}}(P'_2, a, D, \mathcal{K}'_{\mathrm{nlr}}(P'_2))$$
  The reason is that $(P'_1, P'_2) \in \mathcal{B}$ and $\mathcal{B}$ is a Markovian bisimulation.

  - For each computation of P'_1 locally reducible with respect to $K_1 \subseteq \mathcal{K}''_{\mathrm{lr}}(P'_1)$ there exists a computation of P'_2 locally reducible with respect to $K_2 \subseteq \mathcal{K}''_{\mathrm{lr}}(P'_2)$ such that for all $D \in \mathbb{P}/\mathcal{B}$:
  $$\mathit{pbtm}_{\mathrm{g}}(P'_1, D, K_1) \; = \; \mathit{pbtm}_{\mathrm{g}}(P'_2, D, K_2)$$
  and vice versa. The reason is that, since $(P'_1, P'_2) \in \mathcal{B}$ and $\mathcal{B}$ is a Markovian bisimulation, for each tree of computations from P'_1 (resp. P'_2) to D locally reducible with respect to K_1 there exists a tree of computations from P'_2 (resp. P'_1) to D locally reducible with respect to K_2, such that all corresponding states traversed by the computations in the two trees form pairs contained in $\mathcal{B}$. Therefore, the two trees contribute to $\mathit{pbtm}_{\mathrm{g}}$ with the same sum of $\mathit{probtime}_{\mathrm{g}}$ measures.

- Secondly, we demonstrate that $P_1 \sim_{\mathrm{MB}} P_2$ implies $P_1 \simeq_{\mathrm{MB,g}} P_2$. Since we have proved that $\sim_{\mathrm{MB}} \subseteq \approx_{\mathrm{MB,g}}$, the equivalence classes of $\approx_{\mathrm{MB,g}}$ are unions of equivalence classes of $\sim_{\mathrm{MB}}$. Thus, if $P_1 \sim_{\mathrm{MB}} P_2$ and we take $a \in \mathit{Name}$ and $D \in \mathbb{P}/\approx_{\mathrm{MB,g}}$ with $D = \bigcup_{i \in I} D_i$ and $D_i \in \mathbb{P}/\sim_{\mathrm{MB}}$ for all $i \in I$, we have:
$$\mathit{rate}(P_1, a, D) \; = \; \sum_{i \in I} \mathit{rate}(P_1, a, D_i) \; = \; \sum_{i \in I} \mathit{rate}(P_2, a, D_i) \; = \; \mathit{rate}(P_2, a, D)$$
which means that $P_1 \simeq_{\mathrm{MB,g}} P_2$.

- Thirdly, we show that $P_1 \simeq_{\mathrm{MB,g}} P_2$ implies $P_1 \approx_{\mathrm{MB,g}} P_2$. Whenever $P_1 \simeq_{\mathrm{MB,g}} P_2$, then $P_1 \approx_{\mathrm{MB,g}} P_2$ because the following holds under the factorization $\mathcal{K}_{\mathrm{lr}}(P_1) = \emptyset \uplus \mathcal{K}_{\mathrm{lr}}(P_1)$ and $\mathcal{K}_{\mathrm{lr}}(P_2) = \emptyset \uplus \mathcal{K}_{\mathrm{lr}}(P_2)$:

– For all $a \in \mathit{Name}$ and $D \in \mathbb{P}/\approx_{\mathrm{MB,g}}$:
$$\mathit{rate}_{\mathrm{g}}(P_1, a, D, \mathcal{K}'_{\mathrm{nlr}}(P_1)) \;=\; \mathit{rate}(P_1, a, D) \;=\; \mathit{rate}(P_2, a, D) \;=\; \mathit{rate}_{\mathrm{g}}(P_2, a, D, \mathcal{K}'_{\mathrm{nlr}}(P_2))$$
The reason is that $P_1 \simeq_{\mathrm{MB,g}} P_2$.

– For each computation of P_1 locally reducible with respect to $K_1 \subseteq \mathcal{K}''_{\mathrm{lr}}(P'_1)$ there exists a computation of P_2 locally reducible with respect to $K_2 \subseteq \mathcal{K}''_{\mathrm{lr}}(P'_2)$ such that for all $D \in \mathbb{P}/\approx_{\mathrm{MB,g}}$:
$$\mathit{pbtm}_{\mathrm{g}}(P_1, D, K_1) \;=\; \mathit{pbtm}_{\mathrm{g}}(P_2, D, K_2)$$
and vice versa. The reason is that, since $P_1 \simeq_{\mathrm{MB,g}} P_2$, both P_1 and P_2 reach in one step the same equivalence classes at the same rates and hence the first step towards D contributes to $\mathit{pbtm}_{\mathrm{g}}$ in the same way for P_1 and P_2. At that point, it is enough to consider among those equivalence classes reached in one step by P_1 and P_2 both D itself (if reachable in one step) and the ones from which it is possible to arrive at D via the continuation of the considered locally reducible computations.

- Fourthly, we prove that the inclusions are strict. For example, we have:
$$<a, \lambda>.<\tau, \mu>.<\tau, \gamma>.\underline{0} \;\not\simeq_{\mathrm{MB}}\; <a, \lambda>.<\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0}$$
$$<a, \lambda>.<\tau, \mu>.<\tau, \gamma>.\underline{0} \;\simeq_{\mathrm{MB,g}}\; <a, \lambda>.<\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0}$$
and:
$$<\tau, \mu>.<\tau, \gamma>.\underline{0} \;\not\simeq_{\mathrm{MB,g}}\; <\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0}$$
$$<\tau, \mu>.<\tau, \gamma>.\underline{0} \;\approx_{\mathrm{MB,g}}\; <\tau, \tfrac{\mu \cdot \gamma}{\mu + \gamma}>.\underline{0}$$

- Finally, the fact that $P_1 \simeq_{\mathrm{MB,g}} P_2$ iff $P_1 \approx_{\mathrm{MB,g}} P_2$ when $P_1, P_2 \in \mathbb{P}$ have no locally reducible computations stems immediately from the definitions of $\simeq_{\mathrm{MB,g}}$ and $\approx_{\mathrm{MB,g}}$. ■

Proof of Thm. 4.14 (p. 21). Let $P_1, P_2 \in \mathbb{P}$ be such that $P_1 \simeq_{\mathrm{MB,g}} P_2$:

1. Let $P_1, P_2 \in \mathbb{P}_{\mathrm{seq}}$. By virtue of Prop. 4.12, from $P_1 \simeq_{\mathrm{MB,g}} P_2$ it follows that $P_1 \approx_{\mathrm{MB,g}} P_2$ and hence P_1 and P_2 belong to the same equivalence class D with respect to $\approx_{\mathrm{MB,g}}$. Given $<a, \lambda> \in \mathit{Act}$, for all $a' \in \mathit{Name}$ and $D' \in \mathbb{P}/\approx_{\mathrm{MB,g}}$ we have that:
$$\mathit{rate}(<a, \lambda>.P_1, a', D') \;=\; \mathit{rate}(<a, \lambda>.P_2, a', D') \;=\; \begin{cases} \lambda & \text{if } a' = a \wedge D' = D \\ 0 & \text{if } a' \neq a \vee D' \neq D \end{cases}$$
Therefore $<a, \lambda>.P_1 \simeq_{\mathrm{MB,g}} <a, \lambda>.P_2$.
2. Let $P_1, P_2 \in \mathbb{P}_{\mathrm{seq}}$. Given $P \in \mathbb{P}_{\mathrm{seq}}$, for all $a \in \mathit{Name}$ and $D \in \mathbb{P}/\approx_{\mathrm{MB,g}}$ we have that:
$$\begin{aligned} \mathit{rate}(P_1 + P, a, D) \;&=\; \mathit{rate}(P_1, a, D) + \mathit{rate}(P, a, D) \;= \\ &=\; \mathit{rate}(P_2, a, D) + \mathit{rate}(P, a, D) \;=\; \mathit{rate}(P_2 + P, a, D) \end{aligned}$$
because $P_1 \simeq_{\mathrm{MB,g}} P_2$. Therefore $P_1 + P \simeq_{\mathrm{MB,g}} P_2 + P$ and $P + P_1 \simeq_{\mathrm{MB,g}} P + P_2$.
3. Given $H \subseteq \mathit{Name}_{\mathrm{v}}$, for all $a \in \mathit{Name}$ and $D \in \mathbb{P}/\approx_{\mathrm{MB,g}}$ there are two cases:
   - If D does not contain any term of the form P/H, then:
   $$\mathit{rate}(P_1/H, a, D) \;=\; 0 \;=\; \mathit{rate}(P_2/H, a, D)$$
   - If $D = [P/H]_{\approx_{\mathrm{MB,g}}}$, then we can exploit the congruence property of $\approx_{\mathrm{MB,g}}$ with respect to the hiding operator as established by Prop. 4.10 in order to express D as $\bigcup_{P' \in D'} [P']_{\approx_{\mathrm{MB,g}}}/H$, where D' is a maximal set including P of process terms that are pairwise not related by $\approx_{\mathrm{MB,g}}$, such that $P'/H \approx_{\mathrm{MB,g}} P/H$ for all $P' \in D'$. As a consequence:
   $$\mathit{rate}(P_1/H, a, D) \;=\; \mathit{rate}(P_2/H, a, D)$$
   because for $i = 1, 2$ it holds that:
   $$\mathit{rate}(P_i/H, a, D) \;=\; \begin{cases} 0 & \text{if } a \in H \\ \sum\limits_{P' \in D'} \mathit{rate}(P_i, a, [P']_{\approx_{\mathrm{MB,g}}}) & \text{if } a \notin H \cup \{\tau\} \\ \sum\limits_{P' \in D'} \sum\limits_{b \in H \cup \{\tau\}} \mathit{rate}(P_h, b, [P']_{\approx_{\mathrm{MB,g}}}) & \text{if } a = \tau \end{cases}$$
   and $P_1 \simeq_{\mathrm{MB,g}} P_2$.

   Therefore $P_1/H \simeq_{\mathrm{MB,g}} P_2/H$.

4. Given $S \subseteq \mathit{Name}_{\mathrm{v}}$ and $P \in \mathbb{P}$, for all $a \in \mathit{Name}$ and $D \in \mathbb{P}/\approx_{\mathrm{MB,g}}$ there are two cases:
   - If D does not contain any term of the form $\bar{P} \,\|_S\, \hat{P}$, then:
     $$\mathit{rate}(P_1 \,\|_S\, P, a, D) \;=\; 0 \;=\; \mathit{rate}(P_2 \,\|_S\, P, a, D)$$
   - If $D = [\bar{P} \,\|_S\, \hat{P}]_{\approx_{\mathrm{MB,g}}}$, then we can exploit the congruence property of $\approx_{\mathrm{MB,g}}$ with respect to the parallel composition operator as established by Prop. 4.10 in order to express D as $\bigcup_{P' \in D'} [P']_{\approx_{\mathrm{MB,g}}} \,\|_S\, \hat{P}$, where D' is a maximal set including $\bar{P}$ of process terms that are pairwise not related by $\approx_{\mathrm{MB,g}}$, such that $P' \,\|_S\, \hat{P} \approx_{\mathrm{MB,g}} \bar{P} \,\|_S\, \hat{P}$ for all $P' \in D'$. As a consequence:
     $$\mathit{rate}(P_1 \,\|_S\, P, a, D) \;=\; \mathit{rate}(P_2 \,\|_S\, P, a, D)$$
     because for $i = 1, 2$ it holds that $\mathit{rate}(P_i \,\|_S\, P, a, D)$ is equal to:
     $$\begin{cases} \sum\limits_{P' \in D'} \mathit{rate}(P_i, a, [P']_{\approx_{\mathrm{MB,g}}}) \otimes \mathit{rate}(P, a, \{\hat{P}\}) & \text{if } a \in S \\ \sum\limits_{P' \in D'} \mathit{rate}(P_i, a, [P']_{\approx_{\mathrm{MB,g}}}) + \mathit{rate}(P, a, \{\hat{P}\}) & \text{if } a \notin S \wedge P_i \in \bigcup\limits_{P' \in D'} [P']_{\approx_{\mathrm{MB,g}}} \wedge P \equiv \hat{P} \\ \sum\limits_{P' \in D'} \mathit{rate}(P_i, a, [P']_{\approx_{\mathrm{MB,g}}}) & \text{if } a \notin S \wedge P_i \notin \bigcup\limits_{P' \in D'} [P']_{\approx_{\mathrm{MB,g}}} \wedge P \equiv \hat{P} \\ \mathit{rate}(P, a, \{\hat{P}\}) & \text{if } a \notin S \wedge P_i \in \bigcup\limits_{P' \in D'} [P']_{\approx_{\mathrm{MB,g}}} \wedge P \not\equiv \hat{P} \\ 0 & \text{if } a \notin S \wedge P_i \notin \bigcup\limits_{P' \in D'} [P']_{\approx_{\mathrm{MB,g}}} \wedge P \not\equiv \hat{P} \end{cases}$$
     and $P_1 \simeq_{\mathrm{MB,g}} P_2$.

   Therefore $P_1 \,\|_S\, P \simeq_{\mathrm{MB,g}} P_2 \,\|_S\, P$ and $P \,\|_S\, P_1 \simeq_{\mathrm{MB,g}} P \,\|_S\, P_2$. ■

Proof of Thm. 4.15 (p. 21). Let $P_1, P_2 \in \mathbb{P}_{\mathrm{seq}}$. The proof is divided into two parts:

- $\Rightarrow$ If $P_1 \simeq_{\mathrm{MB,g}} P_2$, then by virtue of Thm. 4.14 it follows that $P_1 + P \simeq_{\mathrm{MB,g}} P_2 + P$ for all $P \in \mathbb{P}_{\mathrm{seq}}$. Due to Prop. 4.12, this implies that $P_1 + P \approx_{\mathrm{MB,g}} P_2 + P$ for all $P \in \mathbb{P}_{\mathrm{seq}}$.
- $\Leftarrow$ Suppose that $P_1 + P \approx_{\mathrm{MB,g}} P_2 + P$ for all $P \in \mathbb{P}_{\mathrm{seq}}$. Since it is possible to find $\bar{P} \in \mathbb{P}_{\mathrm{seq}}$ such that neither $P_1 + \bar{P}$ nor $P_2 + \bar{P}$ has locally reducible computations, from $P_1 + \bar{P} \approx_{\mathrm{MB,g}} P_2 + \bar{P}$ it follows that $P_1 + \bar{P} \simeq_{\mathrm{MB,g}} P_2 + \bar{P}$ because $\simeq_{\mathrm{MB,g}}$ and $\approx_{\mathrm{MB,g}}$ coincide over the set of process terms that have no locally reducible computations as established by Prop. 4.12. Since for all $a \in \mathit{Name}$ and $D \in \mathbb{P}/\approx_{\mathrm{MB,g}}$ it then holds that:
  $$\begin{aligned} \mathit{rate}(P_1, a, D) \;&=\; \mathit{rate}(P_1 + \bar{P}, a, D) - \mathit{rate}(\bar{P}, a, D) \;= \\ &=\; \mathit{rate}(P_2 + \bar{P}, a, D) - \mathit{rate}(\bar{P}, a, D) \;=\; \mathit{rate}(P_2, a, D) \end{aligned}$$
  we have that $P_1 \simeq_{\mathrm{MB,g}} P_2$. ■

Proof of Prop. 4.18 (p. 21). Let $\mathcal{B}$ be a generalized weak Markovian bisimulation up to $\approx_{\mathrm{MB,g}}$. We first show that $(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+$ is a generalized weak Markovian bisimulation by proving by induction on $n \in \mathbb{N}_{\geq 1}$ that, whenever $(P_1, P_2) \in (\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^n$, then $\mathcal{K}_{\mathrm{lr}}(P_1) = \mathcal{K}'_{\mathrm{lr}}(P_1) \uplus \mathcal{K}''_{\mathrm{lr}}(P_1)$ and $\mathcal{K}_{\mathrm{lr}}(P_2) = \mathcal{K}'_{\mathrm{lr}}(P_2) \uplus \mathcal{K}''_{\mathrm{lr}}(P_2)$ such that:

- For all $a \in \mathit{Name}$ and $D \in \mathbb{P}/(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+$:
  $$\mathit{rate}(P_1, a, D, \mathcal{K}'_{\mathrm{nlr}}(P_1)) \;=\; \mathit{rate}(P_2, a, D, \mathcal{K}'_{\mathrm{nlr}}(P_2))$$
- For each computation of P_1 locally reducible with respect to $K_1 \subseteq \mathcal{K}''_{\mathrm{lr}}(P_1)$ there exists a computation of P_2 locally reducible with respect to $K_2 \subseteq \mathcal{K}''_{\mathrm{lr}}(P_1)$ such that for all $D \in \mathbb{P}/(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+$:
  $$\mathit{pbtm}_{\mathrm{g}}(P_1, D, K_1) \;=\; \mathit{pbtm}_{\mathrm{g}}(P_2, D, K_2)$$

Let $(P_1, P_2) \in (\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^n$:

- If $n = 1$, then $(P_1, P_2) \in \mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}}$. There are two cases:
  - If $(P_1, P_2) \in \mathcal{B} \cup \mathcal{B}^{-1}$, then the result immediately follows from the fact that $\mathcal{B}$ is a generalized weak Markovian bisimulation up to $\approx_{\mathrm{MB,g}}$.

– If $(P_1, P_2) \in \approx_{\rm MB,g}$, then the result stems from the fact that $\approx_{\rm MB,g} \subseteq (\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB,g})^+$ and hence each equivalence class of $(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB,g})^+$ is the union of some equivalence classes of $\approx_{\rm MB,g}$.

- Let $n > 1$ and suppose that the result holds for all $(Q_1, Q_2) \in (\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB,g})^{n-1}$. From $(P_1, P_2) \in (\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB,g})^n$, we derive that there exists $P \in \mathbb{P}$ such that $(P_1, P) \in (\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB,g})^{n-1}$ and $(P, P_2) \in \mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB,g}$. Then the result holds both for the pair (P_1, P) – by the induction hypothesis – and for the pair (P, P_2) – by reasoning like in the case $n = 1$. As a consequence, the result follows for the pair (P_1, P_2) by transitivity of $\mathit{rate}_{\rm g}$ equality and $\mathit{pbtm}_{\rm g}$ equality.

Since we have proved that $(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB,g})^+$ is a generalized weak Markovian bisimulation, it holds that $(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB,g})^+ \subseteq \approx_{\rm MB,g}$. On the other hand, $\mathcal{B} \subseteq (\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB,g})^+$. Therefore, $\mathcal{B} \subseteq \approx_{\rm MB,g}$ by transitivity of set inclusion, i.e., $(P_1, P_2) \in \mathcal{B}$ implies $P_1 \approx_{\rm MB,g} P_2$ for all $P_1, P_2 \in \mathbb{P}$. We also note that $(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB,g})^+ = \approx_{\rm MB,g}$. ■

Proof of Thm. 4.19 (p. 21). Without loss of generality, we assume for simplicity that the two sequential process terms $P_1, P_2 \in \mathcal{PL}$ such that $P_1 \simeq_{\rm MB,g} P_2$ contain free occurrences of a single process variable $X \in \mathit{Var}$. Consider the binary relation:

$$\mathcal{B} = \{(P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}) \mid P \in \mathcal{PL} \text{ containing free occurrences of } X \text{ at most}\}$$

which is formed by pairs each of which is composed of two process terms such that neither of them has a locally reducible computation, or both of them have. In fact, e.g., it is not possible that $P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ has no locally reducible computations while $P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}$ has a locally reducible computation because:

- If P is not a process variable, then the actions enabled by $P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ and the actions enabled by $P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}$ coincide with the actions enabled by P.

- If P is a process variable, which must be X, then $P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ is equal to $\mathrm{rec}\, X : P_1$ and $P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}$ is equal to $\mathrm{rec}\, X : P_2$. The two resulting process terms are isomorphic to $P_1\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ and $P_2\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}$, respectively, with $P_1\{\mathrm{rec}\, X : P_1 \hookrightarrow X\} \simeq_{\rm MB,g} P_2\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}$ because $P_1 \simeq_{\rm MB,g} P_2$.

Similar to [27], we show that $\mathcal{B}$ has a property stronger than being a generalized weak Markovian bisimulation up to $\approx_{\rm MB,g}$: for each $P \in \mathcal{PL}$ containing free occurrences of X at most, it holds that for all action names $a \in \mathit{Name}$ and equivalence classes $D \in \mathcal{PL}/(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\rm MB,g})^+$:

$$\mathit{rate}(P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, a, D) \;\leq\; \mathit{rate}(P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}, a, D)$$

(like in [15], $\geq$ can be established between the two *rate* values with a symmetric argument, from which it can be concluded that the two *rate* values coincide).

If $\mathit{rate}(P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, a, D) = 0$, then the property trivially holds, otherwise we proceed by induction on the maximum depth $d \in \mathbb{N}_{\geq 1}$ of the inferences of the transitions from $P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ to D labeled with a:

- If $d = 1$, then only the semantic rule for the action prefix operator has been applied and hence P must be of the form $<a, \lambda>.P'$ (notice that it cannot be P equal to X because in that case $P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ would be equal to $\mathrm{rec}\, X : P_1$, which would contradict $d = 1$). Thus, for $i = 1, 2$ we have that $P\{\mathrm{rec}\, X : P_i \hookrightarrow X\}$ is of the form $<a, \lambda>.(P'\{\mathrm{rec}\, X : P_i \hookrightarrow X\}$. Since P' contains free occurrences of X at most, $(P'\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, P'\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}) \in \mathcal{B}$ and hence both process terms belong to D. Thus $\mathit{rate}(P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, a, D) = \lambda = \mathit{rate}(P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}, a, D)$.

- Let $d > 1$ and suppose that the property holds for all triples composed of a pair of process terms in $\mathcal{B}$, an equivalence class D', and an action name a' such that there are transitions from the first process term of the pair to D' labeled with a' and the maximum depth of their inferences is at most $d - 1$. We have the following cases:

- If P is of the form $P' + P''$, then for $i = 1, 2$ we have that $P\{\mathrm{rec}\, X : P_i \hookrightarrow X\}$ is of the form $P'\{\mathrm{rec}\, X : P_i \hookrightarrow X\} + P''\{\mathrm{rec}\, X : P_i \hookrightarrow X\}$ and hence $\mathit{rate}(P\{\mathrm{rec}\, X : P_i \hookrightarrow X\}, a, D) = \mathit{rate}(P'\{\mathrm{rec}\, X : P_i \hookrightarrow X\}, a, D) + \mathit{rate}(P''\{\mathrm{rec}\, X : P_i \hookrightarrow X\}, a, D)$. In this case, the semantic rules for the alternative composition operator are applied first and hence the transitions from $P'\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ and $P''\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ to D labeled with a are considered (their inferences have maximum depth $d - 1$). If there are no such transitions from $P'\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$, then $\mathit{rate}(P'\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, a, D) = 0$, otherwise – since P' contains free occurrences of X at most – from the induction hypothesis it follows that $\mathit{rate}(P'\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, a, D) \leq \mathit{rate}(P'\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}, a, D)$. Using a similar argument, we have that $\mathit{rate}(P''\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, a, D) = 0$ or by the induction hypothesis $\mathit{rate}(P''\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, a, D) \leq \mathit{rate}(P''\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}, a, D)$. Thus $\mathit{rate}(P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, a, D) \leq \mathit{rate}(P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}, a, D)$.
- If P is a process variable, which must be X, then for $i = 1, 2$ we have that $P\{\mathrm{rec}\, X : P_i \hookrightarrow X\}$ is equal to $\mathrm{rec}\, X : P_i$, which in turn is isomorphic to $P_i\{\mathrm{rec}\, X : P_i \hookrightarrow X\}$ and hence $\mathit{rate}(\mathrm{rec}\, X : P_i, a, D) = \mathit{rate}(P_i\{\mathrm{rec}\, X : P_i \hookrightarrow X\}, a, D)$. In this case, the semantic rule for recursion is applied first and hence the transitions from $P_1\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ to D labeled with a are considered (their inferences have maximum depth $d - 1$). Since P_1 contains free occurrences of X at most, from the induction hypothesis it follows that $\mathit{rate}(P_1\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, a, D) \leq \mathit{rate}(P_1\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}, a, D)$, with $\mathit{rate}(P_1\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}, a, D) = \mathit{rate}(P_2\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}, a, D)$ because $P_1 \simeq_{\mathrm{MB,g}} P_2$. Thus $\mathit{rate}(\mathrm{rec}\, X : P_1, a, D) \leq \mathit{rate}(\mathrm{rec}\, X : P_2, a, D)$.
- If P is of the form $\mathrm{rec}\, Y : P'$, then there are two subcases:
  - ∗ If $Y = X$, then P contains no free occurrences of X. Therefore, for $i = 1, 2$ we have that $P\{\mathrm{rec}\, X : P_i \hookrightarrow X\}$ is equal to P and hence $\mathit{rate}(P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, a, D) = \mathit{rate}(P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}, a, D)$.
  - ∗ If $Y \neq X$, then for $i = 1, 2$ we have that $P\{\mathrm{rec}\, X : P_i \hookrightarrow X\}$ is isomorphic to $P'\{\mathrm{rec}\, Y : P' \hookrightarrow Y\}\{\mathrm{rec}\, X : P_i \hookrightarrow X\}$ and hence $\mathit{rate}(P\{\mathrm{rec}\, X : P_i \hookrightarrow X\}, a, D) = \mathit{rate}(P'\{\mathrm{rec}\, Y : P' \hookrightarrow Y\}\{\mathrm{rec}\, X : P_i \hookrightarrow X\}, a, D)$. In this case, the semantic rule for recursion is applied first and hence the transitions from $P'\{\mathrm{rec}\, Y : P' \hookrightarrow Y\}\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ to D labeled with a are considered (their inferences have maximum depth $d-1$). Since $P'\{\mathrm{rec}\, Y : P' \hookrightarrow Y\}$ contains free occurrences of X at most, from the induction hypothesis it follows that $\mathit{rate}(P'\{\mathrm{rec}\, Y : P' \hookrightarrow Y\}\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, a, D) \leq \mathit{rate}(P'\{\mathrm{rec}\, Y : P' \hookrightarrow Y\}\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}, a, D)$. Thus $\mathit{rate}(P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, a, D) \leq \mathit{rate}(P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}, a, D)$.
- If P is of the form P'/H, then for $i = 1, 2$ we have that $P\{\mathrm{rec}\, X : P_i \hookrightarrow X\}$ is of the form $(P'\{\mathrm{rec}\, X : P_i \hookrightarrow X\})/H$. In this case, the semantic rules for the hiding operator are applied first and hence the transitions from $P'\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ to $[Q']_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+}$ labeled with a' are considered (their inferences have maximum depth $d - 1$), where $Q'/H \in D$ and $a' = a$ for $a \notin H \cup \{\tau\}$, $a' \in H \cup \{\tau\}$ for $a = \tau$. Since $\approx_{\mathrm{MB,g}}$ is a congruence with respect to the hiding operator by virtue of Prop. 4.10, and $\mathcal{B}$ can be shown to be a congruence with respect to the hiding operator, the equivalence class D can be expressed as $\bigcup_{Q'' \in D'} [Q'']_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+}/H$, where D' is a maximal set including Q' of process terms that are pairwise not related by $(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+$, such that $(Q''/H, Q'/H) \in (\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+$ for all $Q'' \in D'$. As a consequence, for $i = 1, 2$ we have that $\mathit{rate}((P'\{\mathrm{rec}\, X : P_i \hookrightarrow X\})/H, a, D)$ is equal to:
  - ∗ 0 if $a \in H$.
  - ∗ $\sum_{Q'' \in D'} \mathit{rate}(P'\{\mathrm{rec}\, X{:}P_i \hookrightarrow X\}, a, [Q'']_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+})$ if $a \notin H \cup \{\tau\}$.
  - ∗ $\sum_{Q'' \in D'} \sum_{a' \in H \cup \{\tau\}} \mathit{rate}(P'\{\mathrm{rec}\, X{:}P_i \hookrightarrow X\}, a', [Q'']_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+})$ if $a = \tau$.

  For each of the rate summands giving rise to the P_1-related value $\mathit{rate}((P'\{\mathrm{rec}\, X : P_1 \hookrightarrow X\})/H, a, D)$, if there are no transitions contributing to the rate summand, then the summand is equal to 0, otherwise – since P' contains free occurrences of X at most – from the induction hypothesis it follows that the summand is not greater than the corresponding rate summand for P_2. Therefore, we conclude that $\mathit{rate}((P'\{\mathrm{rec}X : P_1 \hookrightarrow X\})/H, a, D) \leq \mathit{rate}((P'\{\mathrm{rec}X : P_2 \hookrightarrow X\})/H, a, D)$.

– If P is of the form $P' \parallel_S P''$, then for $i = 1, 2$ we have that $P\{\mathrm{rec}\, X : P_i \hookrightarrow X\}$ is of the form $P'\{\mathrm{rec}\, X : P_i \hookrightarrow X\} \parallel_S P''\{\mathrm{rec}\, X : P_i \hookrightarrow X\}$. In this case, the semantic rules for the parallel composition operator are applied first and hence the transitions from $P'\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ to $[Q']_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+}$ and from $P''\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ to $[Q'']_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+}$ labeled with a are considered (their inferences have maximum depth $d - 1$), where $Q' \parallel_S Q'' \in D$. Since $\approx_{\mathrm{MB,g}}$ is a congruence with respect to the parallel composition operator by virtue of Prop. 4.10, and $\mathcal{B}$ can be shown to be a congruence with respect to the parallel composition operator, the equivalence class D can be expressed as $\bigcup_{Q_1 \in D'_1} \bigcup_{Q_2 \in D'_2} [Q_1]_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+} \parallel_S [Q_2]_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+}$, where D'_1 (resp. D'_2) is a maximal set including Q' (resp. Q'') of process terms that are pairwise not related by $(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+$, such that $(Q_1 \parallel_S Q'', Q' \parallel_S Q'') \in (\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+$ for all $Q_1 \in D'_1$ (resp. $(Q' \parallel_S Q_2, Q' \parallel_S Q'') \in (\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+$ for all $Q_2 \in D'_2$). As a consequence, for $i = 1, 2$ we have that $\mathit{rate}((P'\{\mathrm{rec}\, X : P_i \hookrightarrow X\}) \parallel_S (P''\{\mathrm{rec}\, X : P_i \hookrightarrow X\}), a, D)$ is equal to:
  - $\sum_{Q_1 \in D'_1} \mathit{rate}(P'\{\mathrm{rec}\, X : P_i \hookrightarrow X\}, a, [Q_1]_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+}) \otimes \sum_{Q_2 \in D'_2} \mathit{rate}(P''\{\mathrm{rec}\, X : P_i \hookrightarrow X\}, a, [Q_2]_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+})$ if $a \in S$.
  - $\sum_{Q_1 \in D'_1} \mathit{rate}(P'\{\mathrm{rec}\, X : P_i \hookrightarrow X\}, a, [Q_1]_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+}) + \sum_{Q_2 \in D'_2} \mathit{rate}(P''\{\mathrm{rec}\, X : P_i \hookrightarrow X\}, a, [Q_2]_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+})$ if $a \notin S$ and $P'\{\mathrm{rec}\, X : P_i \hookrightarrow X\} \in \bigcup_{Q_1 \in D'_1} [Q_1]_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+}$ and $P''\{\mathrm{rec}\, X : P_i \hookrightarrow X\} \in \bigcup_{Q_2 \in D'_2} [Q_2]_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+}$.
  - $\sum_{Q_1 \in D'_1} \mathit{rate}(P'\{\mathrm{rec}\, X : P_i \hookrightarrow X\}, a, [Q_1]_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+})$ if $a \notin S$ and $P'\{\mathrm{rec}\, X : P_i \hookrightarrow X\} \notin \bigcup_{Q_1 \in D'_1} [Q_1]_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+}$ and $P''\{\mathrm{rec}\, X : P_i \hookrightarrow X\} \in \bigcup_{Q_2 \in D'_2} [Q_2]_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+}$.
  - $\sum_{Q_2 \in D'_2} \mathit{rate}(P''\{\mathrm{rec}\, X : P_i \hookrightarrow X\}, a, [Q_2]_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+})$ if $a \notin S$ and $P'\{\mathrm{rec}\, X : P_i \hookrightarrow X\} \in \bigcup_{Q_1 \in D'_1} [Q_1]_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+}$ and $P''\{\mathrm{rec}\, X : P_i \hookrightarrow X\} \notin \bigcup_{Q_2 \in D'_2} [Q_2]_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+}$.
  - 0 if $a \notin S$ and $P'\{\mathrm{rec}\, X : P_i \hookrightarrow X\} \notin \bigcup_{Q_1 \in D'_1} [Q_1]_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+}$ and $P''\{\mathrm{rec}\, X : P_i \hookrightarrow X\} \notin \bigcup_{Q_2 \in D'_2} [Q_2]_{(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+}$.

  For each of the *rate* summands giving rise to the P_1-related *rate* value $\mathit{rate}((P'\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}) \parallel_S (P''\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}), a, D)$, if there are no transitions contributing to the *rate* summand, then the summand is equal to 0, otherwise – since P' and P'' contain free occurrences of X at most – from the induction hypothesis it follows that the summand is not greater than the corresponding *rate* summand for P_2. Therefore $\mathit{rate}((P'\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}) \parallel_S (P''\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}), a, D) \leq \mathit{rate}((P'\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}) \parallel_S (P''\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}), a, D)$.

From the property of $\mathcal{B}$ that we have proved (and the symmetrical property), it follows that $\mathcal{B}$ is a generalized weak Markovian bisimulation up to $\approx_{\mathrm{MB,g}}$. In fact, if both $P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ and $P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}$ have locally reducible computations, then for each computation of $P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ locally reducible with respect to some K_1 there exists a computation of $P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}$ locally reducible with respect to some K_2 such that for all equivalence classes $D \in \mathcal{PL}/(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+$ it holds that $\mathit{pbtm}_{\mathrm{g}}(P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, D, K_1) = \mathit{pbtm}_{\mathrm{g}}(P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}, D, K_2)$ and vice versa. The reason is that both $P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ and $P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}$ reach in one step the same equivalence classes at the same rates and hence the first step towards D contributes to $\mathit{pbtm}_{\mathrm{g}}$ in the same way for $P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}$ and $P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}$.

Therefore, by virtue of Prop. 4.18 we have that $(\mathcal{B} \cup \mathcal{B}^{-1} \cup \approx_{\mathrm{MB,g}})^+ = \approx_{\mathrm{MB,g}}$ and hence what we have proved is that, for each $P \in \mathcal{PL}$ containing free occurrences of X at most, it holds that for all $a \in \mathit{Name}$ and $D \in \mathcal{PL}/\approx_{\mathrm{MB,g}}$:

$$\mathit{rate}(P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\}, a, D) \;=\; \mathit{rate}(P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}, a, D)$$

This means that $P\{\mathrm{rec}\, X : P_1 \hookrightarrow X\} \simeq_{\mathrm{MB,g}} P\{\mathrm{rec}\, X : P_2 \hookrightarrow X\}$ for all $P \in \mathcal{PL}$ containing free occurrences of X at most. We finally derive $\mathrm{rec}\, X : P_1 \simeq_{\mathrm{MB,g}} \mathrm{rec}\, X : P_2$ by taking P equal to X. ■

Proof of Thm. 4.22 (p. 24). Similar to the proof of Thm. 3.24, we focus on the rewriting rule in Fig. 6. Let us view the original CTMC on the left and the aggregated CTMC on the right as GSMP models in which all the elements have exponentially distributed lifetimes. Moreover, let us consider an intermediate GSMP having the same state space as the aggregated one, whose transitions corresponding to locally reducible

computations in the original GSMP represent elements whose lifetime is hypoexponentially distributed. In particular, let us call z' the state of the intermediate GSMP corresponding to state z. The transitions of z' corresponding to the transitions of z depicted in Fig. 6 represent an element $\mathcal{H}$ following a hypoexponential distribution with two stages having rate μ and γ, respectively. The death of this element causes z' to reach state $s'_{i,j}$ of the intermediate GSMP corresponding to state $s_{i,j}$ with probability $\frac{\mu_i}{\mu} \cdot \frac{\gamma_{i,j}}{\gamma}$.
The insensitivity balance equation for z' with respect to element $\mathcal{H}$ is as follows:

$$\pi[z'] \cdot \left(\frac{1}{\mu} + \frac{1}{\gamma}\right)^{-1} \; = \; \sum_{s'' \in S_{\mathcal{H}}} \pi[s''] \cdot R(s'', z')$$

where $S_{\mathcal{H}}$ is the set of states in which $\mathcal{H}$ is not active such that their transitions lead to the birth of $\mathcal{H}$ in z'. Since $\mathcal{H}$ and the corresponding exponential element in the aggregated GSMP have the same mean $\frac{1}{\mu} + \frac{1}{\gamma}$, and the various states with transitions entering z' lead to the birth in z' of as many different elements due to the absence of non-joining synchronizations, the sum of the insensitivity balance equations for z' with respect to each of its active elements is equivalent to the global balance equation for z:

$$\pi_{\mathrm{r}}[z] \cdot \left(\left(\frac{1}{\mu} + \frac{1}{\gamma}\right)^{-1} + \Lambda\right) \; = \; \sum_{s'' \in S} \pi_{\mathrm{r}}[s''] \cdot R(s'', z)$$

As a consequence, the intermediate GSMP is insensitive due to Matthes' theorem and its stationary behavior coincides with the one of the aggregated GSMP.
Following the scheme of the proof of Thm. 3.24, the insensitivity balance equations of the original GSMP with respect to the exponential elements corresponding to the stages of $\mathcal{H}$ can be transformed into a form equivalent to the insensitivity balance equations of the intermediate GSMP with respect to $\mathcal{H}$. Therefore, also the original GSMP has the same stationary behavior as the aggregated one and it holds that:

$$\pi_{\mathrm{r}}[z] \; = \; \pi_{\mathrm{l}}[s] + \sum_{i \in I} \pi_{\mathrm{l}}[s_i]$$

■

References

[1] L. Aceto and D. Murphy. Timing and causality in process algebra. *Acta Informatica*, 33:317–350, 1996.
[2] M. Ajmone Marsan, G. Balbo, G. Conte, S. Donatelli, and G. Franceschinis. *Modelling with Generalized Stochastic Petri Nets.* John Wiley & Sons, 1995.
[3] A. Aldini, M. Bernardo, and F. Corradini. *A Process Algebraic Approach to Software Architecture Design.* Springer, 2010.
[4] S. Andova and J.C.M. Baeten. Abstraction in probabilistic process algebra. In *Proc. of the 7th Int. Conf. on Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2001)*, volume 2031 of *LNCS*, pages 204–219. Springer, 2001.
[5] F. Baccelli and P. Bremand. The insensitivity balance equations. In *Palm Probabilities and Stationary Processes*, volume 41 of *LNS*, pages 69–73. Springer, 1987.
[6] C. Baier and H. Hermanns. Weak bisimulation for fully probabilistic processes. In *Proc. of the 9th Int. Conf. on Computer Aided Verification (CAV 1997)*, volume 1254 of *LNCS*, pages 119–130. Springer, 1997.
[7] C. Baier, J.-P. Katoen, H. Hermanns, and V. Wolf. Comparative branching-time semantics for Markov chains. *Information and Computation*, 200:149–214, 2005.
[8] M. Bernardo. Non-bisimulation-based Markovian behavioral equivalences. *Journal of Logic and Algebraic Programming*, 72:3–49, 2007.
[9] M. Bernardo. On the expressiveness of Markovian process calculi with durational and durationless actions. In *Proc. of the 1st Int. Symp. on Games, Automata, Logics and Formal Verification (GANDALF 2010)*, volume 25 of *EPTCS*, pages 199–213, 2010.
[10] M. Bernardo. Weak Markovian bisimulation congruences and exact CTMC-level aggregations for sequential processes. In *Proc. of the 6th Int. Symp. on Trustworthy Global Computing (TGC 2011)*, volume 7173 of *LNCS*, pages 89–103. Springer, 2011.
[11] M. Bernardo. Weak Markovian bisimulation congruences and exact CTMC-level aggregations for concurrent processes. In *Proc. of the 10th Int. Workshop on Quantitative Aspects of Programming Languages and Systems (QAPL 2012)*, volume 85 of *EPTCS*, pages 122–136, 2012.
[12] M. Bernardo and A. Aldini. Weak Markovian bisimilarity: Abstracting from prioritized/weighted internal immediate actions. In *Proc. of the 10th Italian Conf. on Theoretical Computer Science (ICTCS 2007)*, pages 39–56. World Scientific, 2007.
[13] M. Bernardo and M. Bravetti. Performance measure sensitive congruences for Markovian process algebras. *Theoretical Computer Science*, 290:117–160, 2003.
[14] M. Bravetti. Revisiting interactive Markov chains. In *Proc. of the 3rd Int. Workshop on Models for Time-Critical Systems (MTCS 2002)*, volume 68(5) of *ENTCS*, pages 1–20. Elsevier, 2002.

[15] M. Bravetti, M. Bernardo, and R. Gorrieri. A note on the congruence proof for recursion in Markovian bisimulation equivalence. In *Proc. of the 6th Int. Workshop on Process Algebra and Performance Modelling (PAPM 1998)*, pages 71–87, 1998.
[16] P. Buchholz. Exact and ordinary lumpability in finite Markov chains. *Journal of Applied Probability*, 31:59–75, 1994.
[17] C. Eisentraut, H. Hermanns, and L. Zhang. On probabilistic automata in continuous time. In *Proc. of the 25th IEEE Symp. on Logic in Computer Science (LICS 2010)*, pages 342–351. IEEE-CS Press, 2010.
[18] R.J. van Glabbeek and W.P. Weijland. Branching time and abstraction in bisimulation semantics. *Journal of the ACM*, 43:555–600, 1996.
[19] W. Henderson and D. Lucic. Aggregation and disaggregation through insensitivity in stochastic Petri nets. *Performance Evaluation*, 17:91–114, 1993.
[20] H. Hermanns. *Interactive Markov Chains.* Springer, 2002. Volume 2428 of LNCS.
[21] H. Hermanns and M. Rettelbach. Syntax, semantics, equivalences, and axioms for MTIPP. In *Proc. of the 2nd Int. Workshop on Process Algebra and Performance Modelling (PAPM 1994)*, pages 71–87. University of Erlangen, Technical Report 27-4, 1994.
[22] J. Hillston. *A Compositional Approach to Performance Modelling.* Cambridge University Press, 1996.
[23] R.A. Howard. *Dynamic Probabilistic Systems.* John Wiley & Sons, 1971.
[24] P.C. Kanellakis and S.A. Smolka. CCS expressions, finite state processes, and three problems of equivalence. *Information and Computation*, 86:43–68, 1990.
[25] J. Markovski and N. Trcka. Lumping Markov chains with silent steps. In *Proc. of the 3rd Int. Conf. on the Quantitative Evaluation of Systems (QEST 2006)*, pages 221–230. IEEE-CS Press, 2006.
[26] K. Matthes. Zur theorie der bedienungsprozesse. In *Proc. of the 3rd Prague Conf. on Information Theory, Statistical Decision Functions and Random Processes*, pages 513–528, 1962.
[27] R. Milner. *Communication and Concurrency.* Prentice Hall, 1989.
[28] F. Moller and C. Tofts. Behavioural abstraction in TCCS. In *Proc. of the 19th Int. Coll. on Automata, Languages and Programming (ICALP 1992)*, volume 623 of *LNCS*, pages 559–570. Springer, 1992.
[29] U. Montanari and V. Sassone. Dynamic congruence vs. progressing bisimulation for CCS. *Fundamenta Informaticae*, 16:171–199, 1992.
[30] R. Paige and R.E. Tarjan. Three partition refinement algorithms. *SIAM Journal on Computing*, 16:973–989, 1987.
[31] M. Rettelbach. Probabilistic branching in Markovian process algebras. *The Computer Journal*, 38:590–599, 1995.
[32] G. Rubino and B. Sericola. Sojourn times in finite Markov processes. *Journal of Applied Probability*, 27:744–756, 1989.
[33] A. Sharma and J.-P. Katoen. Weighted lumpability on Markov chains. In *Proc. of the 8th Int. Ershov Memorial Conf. on Perspectives of Systems Informatics (PSI 2011)*, volume 7162 of *LNCS*, pages 322–339. Springer, 2011.
[34] W.J. Stewart. *Introduction to the Numerical Solution of Markov Chains.* Princeton University Press, 1994.
[35] A. Valmari and G. Franceschinis. Simple $O(m \log n)$ time Markov chain lumping. In *Proc. of the 16th Int. Conf. on Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2010)*, volume 6015 of *LNCS*, pages 38–52. Springer, 2010.
[36] Wang Yi. CCS + time = an interleaving model for real time systems. In *Proc. of the 18th Int. Coll. on Automata, Languages and Programming (ICALP 1991)*, volume 510 of *LNCS*, pages 217–228. Springer, 1991.