Weak Simulation Quasimetric in a Gossip Scenario

Ruggero Lanotte, Massimo Merro, Simone Tini

Weak simulation quasimetric in a gossip scenario

Ruggero Lanotte[1], Massimo Merro[2], and Simone Tini[1]

[1] Dipartimento di Scienza e Alta Tecnologia, Università dell'Insubria, Como, Italy
[2] Dipartimento di Informatica, Università degli Studi di Verona, Italy

Abstract We propose the notion of weak simulation quasimetric as the quantitative counterpart of weak simulation for probabilistic processes. This is an asymmetric variant of the weak bisimulation metric of Desharnais et al. which maintains most of the properties of the original definition. However, our asymmetric version is particularly suitable to reason on protocols where the systems under consideration are not approximately equivalent. As a main application, we adopt our simulation theory in a simple probabilistic timed process calculus to derive an algebraic theory to evaluate the performances of gossip protocols.

1 Introduction

Behavioural semantics, such as *preorders* and *equivalences*, provide formal instruments to compare the behaviour of *probabilistic systems* [16]. Preorders allow us to determine whether a system can mimic the stepwise behaviour of another system; whereas equivalences require a sort of mutual simulation between two systems. The most prominent examples are the *simulation preorder* and the *bisimulation equivalence* [22,25]. Since probability values usually originate from observations (statistical sampling) or from requirements (probabilistic specification), both preorders and equivalences are only partially satisfactory as they can only say whether a system can mimic another one. Any tiny variation of the probabilistic behaviour of a system will break the preorder (resp. equivalence) without any further information. In practice, many system implementations can only approximate the system specification; thus, the verification of such implementations requires appropriate instruments to measure the quality of the approximation. To this end, *metric semantics* [9,4,6] have been successfully employed to formalise the *behavioural distance* between two systems.

Since metric semantics are inherently symmetric, they can be applied only when dealing with systems which are approximately equivalent. In this paper, we propose the notion of *weak simulation quasimetric* which is the asymmetric counterpart of the *weak bisimulation metric* [10], and the quantitative analogous of the *weak simulation preorder* [2,1]. We use the definition of weak simulation quasimetric to derive a definition of *weak simulation with tolerance* $p \in [0,1]$ between two probabilistic systems; being 0 and 1 the minimum and the maximum tolerance, respectively. Thus, we will write $S \sqsubseteq_p S'$ if the system S' is able to simulate the stepwise behaviour of the system S with a tolerance (or distance) p: for $p = 0$ the two systems are weakly similar in standard manner, while for $p = 1$ they are potentially unrelated.

Our weak simulation with tolerance is suitable for compositional reasonings. The compositionality of a behavioural semantics with respect to the parallel operator is fundamental when reasoning on large-scale systems. Several quantitative analogous of the well-known notions of precongruence (and congruence) have been proposed [10,12] to ensure that systems are *approximately inter-substitutable*. We prove that weak simulation with tolerance matches one of the strongest one, namely *non-expansiveness*:

$$S_1 \sqsubseteq_{p_1} S_1' \text{ and } S_2 \sqsubseteq_{p_2} S_2' \text{ entails } S_1 \mid S_2 \sqsubseteq_{p_1+p_2} S_1' \mid S_2' \ .$$

As (non-trivial) case study, we apply our simulation theory to study and estimate the performance of *gossip networks* for Wireless Sensor Networks (WSNs).

Gossip protocols [17] rely on algorithms to deliver data packets in a network from a source to a destination. They address some critical problems of *flooding*, where each node that receives a message propagates it to all its neighbours by broadcast. The goal of gossip protocols is to reduce the number of retransmissions by making some of the nodes discard the message instead of forwarding it. Gossip protocols exhibit both *nondeterministic* and *probabilistic* behaviour. Nondeterminism arises as they deal with distributed networks in which the activities of individual nodes occur nondeterministically. As to the probabilistic behaviour, nodes are required to forward packets with a pre-specified gossip probability p_{gsp}. When a node receives a message, rather than immediately retransmitting it as in flooding, it relies on the probability p_{gsp} to determine whether or not to retransmit. The main benefit is that when p_{gsp} is sufficiently large, the entire network receives the broadcast message with very high probability, even though only a nondeterministic subset of nodes has forwarded the message.

In this paper, we rely on our simulation with tolerance to develop an *algebraic theory* for a simple probabilistic distributed timed calculus [19,23,24,5] which is particularly suitable to represent gossip networks. Our algebraic theory is also compositional as it allows us to join, and sometime merge, the tolerances of different sub-networks with different behaviours. Last but not least, our algebraic theory can be easily *mechanised*. In this extended abstract proofs are omitted.

2 A probabilistic timed process calculus

In Table 1 we define the syntax of the *Probabilistic Timed Calculus for Wireless Systems* [19], `pTCWS`, in a two-level structure, a lower one for *processes*, ranged over by letters P, Q and R, and an upper one for *networks*, ranged over by letters M, N, and O. We use letters $m, n, \ldots$ for logical names; greek symbols $\mu, \nu, \nu_1, \ldots$ for *sets of names*; x, y, z for *variables*; u for *values*, and v and w for *closed values*, i.e. values that do not contain variables. Then, we use p_i for probability weights, hence $p_i \in [0, 1]$.

A network in `pTCWS` is a (possibly empty) collection of nodes (which represent devices) running in parallel and using a common radio channel to communicate with each other. Nodes are unique; i.e. a node n can occur in a network only once. All nodes are assumed to have the same transmission range. The communication paradigm is *local broadcast*; only nodes located in the range of the

Networks:

M, N	$::= \mathbf{0}$	empty network
	$\mid M_1 \mid M_2$	parallel composition
	$\mid n[P]^{\nu}$	node
	$\mid \mathsf{Dead}$	stucking network

Processes:

P, Q	$::= \mathsf{nil}$	termination
	$\mid !\langle u \rangle.C$	broadcast
	$\mid \lfloor ?(x).C \rfloor D$	receiver with timeout
	$\mid \tau.C$	internal
	$\mid \sigma.C$	sleep
	$\mid X$	process variable
	$\mid \mathsf{fix}\, X.P$	recursion

Probabilistic Choice:

C, D	$::= \bigoplus_{i \in I} p_i{:}P_i$	

Table 1. Syntax

transmitter may receive data. We write $n[P]^{\nu}$ for a node named n (the device network address) executing the sequential process P. The set ν contains (the names of) the neighbours of n. Said in other words, ν contains all nodes laying in the transmission cell of n (except n). In this manner, we model the network topology. Our wireless networks have a fixed topology. Moreover, nodes cannot be created or destroyed. Finally, we write Dead to denote a deadlocked network which prevents the execution of parallel components. This is a fictitious network which is introduced for technical convenience (see Definition 9) and not for specifying gossip protocols.

Processes are sequential and live inside the nodes. The symbol nil denotes terminated processes. The sender process $!\langle v \rangle.C$ broadcasts the value v, the continuation being C. The process $\lfloor ?(x).C \rfloor D$ denotes a receiver with timeout. Intuitively, this process either receives a value v, in the current time interval, and then continues as C where the variable x is instantiated with v, or it idles for one time unit, and then continues as D. The process $\tau.C$ performs an internal action and then continues as C. The process $\sigma.C$ models sleeping for one time unit. In processes of the form $\sigma.D$ and $\lfloor ?(x).C \rfloor D$ the occurrence of D is said to be *time-guarded*. The process $\mathsf{fix}\, X.P$ denotes *time-guarded recursion*, as all occurrences of the process variable X may only occur time-guarded in P. With an abuse of notation, we will write $?(x).C$ as an abbreviation for $\mathsf{fix}\, X.\lfloor ?(x).C \rfloor (1{:}X)$, where the process variable X does not occur in C.

The construct $\bigoplus_{i \in I} p_i{:}P_i$ denotes *probabilistic choice*, where I is a *finite*, *non-empty* set of indexes, and $p_i \in (0, 1]$ denotes the probability to execute the process P_i, with $\sum_{i \in I} p_i = 1$. Notice that, as in [8], in order to simplify the operational semantics, probabilistic choices occur always underneath prefixing.

In processes of the form $\lfloor ?(x).C \rfloor D$ the variable x is bound in C. Similarly, in process $\mathsf{fix}\, X.P$ the process variable X is bound in P. This gives rise to the standard notions of *free (process) variables* and *bound (process) variables* and

α-conversion. We identify processes and networks up to α-conversion. A process (resp. probabilistic choice) is said to be *closed* if it does not contain free (process) variables. We always work with closed processes (resp. probabilistic choices): the absence of free variables is trivially maintained at run-time. We write $\{^v/_x\}P$ (resp. $\{^v/_x\}C$) for the substitution of the variable x with the value v in the process P (resp. probabilistic choice C). Similarly, we write $\{^P/_X\}Q$ for the substitution of the process variable X with the process P in Q.

We report some *notational conventions*. $\prod_{i\in I} M_i$ denotes the parallel composition of all M_i, for $i \in I$. We write $P_1 \oplus_p P_2$ for the probabilistic process $p{:}P_1 \oplus (1{-}p){:}P_2$. We identify $1{:}P$ with P. We write $!\langle v\rangle$ as an abbreviation for $!\langle v\rangle.1{:}\mathsf{nil}$. For $k > 0$ we write $\sigma^k.P$ as an abbreviation for $\sigma.\ldots\sigma.P$, where prefix σ appears k times. Given a network M, $\mathsf{nds}(M)$ returns the names of M. If $m \in \mathsf{nds}(M)$, the function $\mathsf{ngh}(m, M)$ returns the set of the neighbours of m in M. Thus, for $M = M_1 \mid m[P]^\nu \mid M_2$ it holds that $\mathsf{ngh}(m, M) = \nu$. We write $\mathsf{ngh}(M)$ for $\bigcup_{m\in\mathsf{nds}(M)} \mathsf{ngh}(m, M)$.

Definition 1. The *structural congruence* over *pTCWS*, written $\equiv$, is defined as the smallest equivalence relation over networks, preserved by parallel composition, which is a commutative monoid with respect to parallel composition with neutral element $\mathbf{0}$, and for which $n[\mathsf{fix}\, X.P]^\nu \equiv n[\{^{\mathsf{fix}\, X.P}/_X\}P]^\nu$.

The syntax presented in Table 1 allows us to derive networks which are somehow ill-formed. With the following definition we rule out networks: (i) where nodes can be neighbours of themselves; (ii) with two different nodes with the same name; (iii) with non-symmetric neighbouring relations. Finally, in order to guarantee clock synchronisation among nodes, we require network connectivity.

Definition 2 (Well-formedness). A network M is said to be *well-formed* if (i) whenever $M \equiv M_1 \mid m[P_1]^\nu$ it holds that $m \notin \nu$; (ii) whenever $M \equiv M_1 \mid m_1[P_1]^{\nu_1} \mid m_2[P_2]^{\nu_2}$ it holds that $m_1 \neq m_2$; (iii) whenever $M \equiv N \mid m_1[P_1]^{\nu_1} \mid m_2[P_2]^{\nu_2}$ we have $m_1 \in \nu_2$ iff $m_2 \in \nu_1$; (iv) for all $m, n \in \mathsf{nds}(M)$ there are $m_1, .., m_k \in \mathsf{nds}(M)$, s.t. $m{=}m_1$, $n{=}m_k$, and $m_i \in \mathsf{ngh}(m_{i+1}, M)$ for $1{\leq}\, i{\leq}k{-}1$.

Henceforth, we will always work with well-formed networks.

2.1 Probabilistic labelled transition semantics

Along the lines of [8], we propose an *operational semantics* for pTCWS associating with each network a graph-like structure representing its possible evolutions: we use a generalisation of labelled transition systems that includes probabilities. Below, we report the mathematical machinery for doing that.

Definition 3. A (discrete) *probability sub-distribution* over a *finite* set S is a function $\Delta\colon S \to [0, 1]$ with $\sum_{s\in S} \Delta(s) \in (0, 1]$. We denote $\sum_{s\in S} \Delta(s)$ by $|\Delta|$. The *support* of a probability sub-distribution Δ is given by $\lceil\Delta\rceil = \{s \in S : \Delta(s) > 0\}$. We write $\mathcal{D}_{\mathrm{sub}}(S)$, ranged over Δ, Θ, Φ, for the set of all probability sub-distributions over S with finite support. A probability sub-distribution $\Delta \in \mathcal{D}_{\mathrm{sub}}(S)$ is said to be a *probability distribution* if $\sum_{s\in S} \Delta(s) = 1$. With $\mathcal{D}(S)$ we

denote the set of all probability distributions over S with finite support. For any $s \in S$, the *point (Dirac) distribution at* s, denoted $\overline{s}$, assigns probability 1 to s and 0 to all others elements of S, so that $\lceil \overline{s} \rceil = \{s\}$.

Let I be a finite index such that (i) Δ_i is a sub-distribution in $\mathcal{D}_{\text{sub}}(S)$ for each $i \in I$, and (ii) $p_i \geq 0$ are probabilities such that $\sum_{i \in I} p_i \in (0, 1]$. Then, the probability sub-distribution $\sum_{i \in I} p_i \cdot \Delta_i \in \mathcal{D}_{\text{sub}}(S)$ is defined as:

$$(\sum_{i \in I} p_i \cdot \Delta_i)(s) \stackrel{\text{def}}{=} \sum_{i \in I} p_i \cdot \Delta_i(s)$$

for all $s \in S$. We write a sub-distribution as $p_1 \cdot \Delta_1 + \ldots + p_n \cdot \Delta_n$ when the index set I is $\{1, \ldots, n\}$. Sometimes, with an abuse of notation, in the previous decomposition we admit that the terms Δ_i are not necessarily distinct (for instance $1 \cdot \Delta$ may be rewritten as $p \cdot \Delta + (1-p) \cdot \Delta$, for any $p \in [0, 1]$). In the following, we will often write $\sum_{i \in I} p_i \Delta_i$ instead of $\sum_{i \in I} p_i \cdot \Delta_i$.

Definition 1 and Definition 2 generalise to sub-distributions in $\mathcal{D}_{\text{sub}}(\texttt{pTCWS})$. Given two sub-distributions Δ and Θ, we write $\Delta \equiv \Theta$ if $\Delta([M]_{\equiv}) = \Theta([M]_{\equiv})$ for all equivalence classes $[M]_{\equiv} \subseteq \texttt{pTCWS}$. A sub-distribution $\Delta \in \mathcal{D}_{\text{sub}}(\texttt{pTCWS})$ is said to be well-formed if its support contains only well-formed networks.

We now give the probabilistic generalisation of labelled transition systems.

Definition 4 (Probabilistic LTS). A *probabilistic labelled transition system* (pLTS) is a triple $\langle S, \mathcal{L}, \rightarrow \rangle$ where (i) S is a set of states; (ii) $\mathcal{L}$ is a set of transition labels; (iii) $\rightarrow$ is a labelled transition relation contained in $S \times \mathcal{L} \times \mathcal{D}(S)$.

The operational semantics of pTCWS is given by a particular pLTS $\langle \texttt{pTCWS}, \mathcal{L}, \rightarrow \rangle$, where $\mathcal{L} = \{m!v \triangleright \mu, m?v, \tau, \sigma\}$ contains the labels denoting broadcasting, reception, internal actions and time passing, respectively. The definition of the relations $\stackrel{\lambda}{\longrightarrow}$, for $\lambda \in \mathcal{L}$, is given by the SOS rules in Table 2. Some of these rules use an obvious notation for distributing parallel composition over a sub-distribution: $(\Delta \mid \Theta)(M) = \Delta(M_1) \cdot \Theta(M_2)$ if $M = M_1 \mid M_2$; $(\Delta \mid \Theta)(M) = 0$ otherwise.

Furthermore, the definition of the labelled transition relation relies on a semantic interpretation of (nodes containing) probabilistic processes in terms of probability distributions over networks.

Definition 5. For any probabilistic choice $\bigoplus_{i \in I} p_i{:}P_i$ over a finite index set I, we write $[\![n[\bigoplus_{i \in I} p_i{:}P_i]^{\mu}]\!]$ to denote the probability distribution $\sum_{i \in I} p_i \cdot \overline{n[P_i]^{\mu}}$.

Let us comment on the most significant rules of Table 2. In rule (Snd) a node m broadcasts a message v to its neighbours ν, the continuation being the probability distribution associated to C. In the label $m!v \triangleright \nu$ the set ν denotes the neighbours of m. In rule (Rcv) a node n gets a message v from a neighbour node m, the continuation being the distribution associated to $\{^v/_x\}C$. If no message is received in the current time interval then the node n will continue according to D, as specified in rule (Timeout). Rules (Rcv-**0**) and (RcvEnb) serve to model reception enabling for synchronisation purposes. For instance, rule (RcvEnb) regards nodes n which are not involved in transmissions originating from m. This

(Snd) $\dfrac{-}{m[!\langle v\rangle.C]^\nu \xrightarrow{m!v\triangleright\nu} [\![m[C]^\nu]\!]}$ (Rcv) $\dfrac{m \in \nu}{n[\lfloor ?(x).C \rfloor D]^\nu \xrightarrow{m?v} [\![n[\{^v/_x\}C]^\nu]\!]}$

(Rcv-$\mathbf{0}$) $\dfrac{-}{\mathbf{0} \xrightarrow{m?v} \overline{\mathbf{0}}}$ (RcvEnb) $\dfrac{\neg(m \in \nu \ \wedge\ \mathsf{rcv}(P)) \ \wedge\ m \neq n}{n[P]^\nu \xrightarrow{m?v} \overline{n[P]^\nu}}$

(RcvPar) $\dfrac{M \xrightarrow{m?v} \Delta \quad N \xrightarrow{m?v} \Theta}{M \mid N \xrightarrow{m?v} \Delta \mid \Theta}$ (Bcast) $\dfrac{M \xrightarrow{m!v\triangleright\nu} \Delta \ \ N \xrightarrow{m?v} \Theta \ \ _{\mu:=\nu\setminus\mathsf{nds}(N)}}{M \mid N \xrightarrow{m!v\triangleright\mu} \Delta \mid \Theta}$

(Tau) $\dfrac{-}{m[\tau.C]^\nu \xrightarrow{\tau} [\![m[C]^\nu]\!]}$ (TauPar) $\dfrac{M \xrightarrow{\tau} \Delta \quad N \neq \mathsf{Dead}}{M \mid N \xrightarrow{\tau} \Delta \mid \overline{N}}$

(σ-$\mathbf{0}$) $\dfrac{-}{\mathbf{0} \xrightarrow{\sigma} \overline{\mathbf{0}}}$ (Timeout) $\dfrac{-}{n[\lfloor ?(x).C \rfloor D]^\nu \xrightarrow{\sigma} [\![n[D]^\nu]\!]}$

(σ-nil) $\dfrac{-}{n[\mathsf{nil}]^\nu \xrightarrow{\sigma} \overline{n[\mathsf{nil}]^\nu}}$ (Sleep) $\dfrac{-}{n[\sigma.C]^\nu \xrightarrow{\sigma} [\![n[C]^\nu]\!]}$

(σ-Par) $\dfrac{M \xrightarrow{\sigma} \Delta \quad N \xrightarrow{\sigma} \Theta}{M \mid N \xrightarrow{\sigma} \Delta \mid \Theta}$ (Rec) $\dfrac{n[\{^{\mathsf{fix}\,X.P}/_X\}P]^\nu \xrightarrow{\lambda} \Delta}{n[\mathsf{fix}\,X.P]^\nu \xrightarrow{\lambda} \Delta}$

Table 2. Probabilistic Labelled Transition System

may happen either because the two nodes are out of range (i.e. $m \notin \nu$) or because n is not willing to receive ($\mathsf{rcv}(P)$ is a boolean predicate that returns true if $n[P]^\nu \equiv n[\lfloor ?(x).C \rfloor D]^\nu$, for some x, C, D). In both cases, node n is not affected by the transmission. Rule (Bcast) models broadcast of messages. Note that we loose track of those transmitter's neighbours that are in N. Rule (Sleep) models sleeping for one time unit. Rule (σ-Par) models time synchronisation between parallel components. Rules (Bcast) and (TauPar) have their symmetric counterparts which are not reported in the table. Finally, note that the semantics of the network Dead is different from that of $\mathbf{0}$: the network Dead does not perform any action and does prevent the evolution of any parallel component.

Extensional labelled transition semantics Our focus is on weak similarities, which abstract away non-observable actions, i.e. those actions that cannot be detected by a parallel network. The adjective *extensional* is used to stress that those activities require a contribution of the environment. To this end, we extend Table 2 by the following two rules:

(ShhSnd) $\dfrac{M \xrightarrow{m!v\triangleright\emptyset} \Delta}{M \xrightarrow{\tau} \Delta}$ (ObsSnd) $\dfrac{M \xrightarrow{m!v\triangleright\nu} \Delta \quad \nu \neq \emptyset}{M \xrightarrow{!v\triangleright\nu} \Delta}$

Rule (ShhSnd) models transmissions that cannot be observed because there is no potential receiver outside the network M. Rule (ObsSnd) models transmissions that can be observed by those nodes of the environment contained in ν. Notice

that the name of the transmitter is removed from the label. This is motivated by the fact that receiver nodes do not have a direct manner to observe the identity of the transmitter. On the other hand, a network M performing the action $m?v$ can be observed by an external node m which transmits the value v to an appropriate set of nodes in M. Notice that the action $!v \triangleright \nu$ does not propagate over parallel components (there is no rule for that). As a consequence, the Rule (ObsSnd) can only be applied to the whole network, never in a sub-network.

In the rest of the paper, the metavariable α will range over the following four kinds of actions: $!v \triangleright \nu$, $m?v$, σ, τ. They denote anonymous broadcast to specific nodes, message reception, time passing, and internal activities, respectively.

3 Weak simulation up to tolerance

In this section, we introduce *weak simulation quasimetrics* as an instrument to derive a notion of approximate simulation between networks. Our goal is to define a family of relations $\sqsubseteq_p$ over networks, with $p \in [0,1]$, to formalise the concept of *simulation with a tolerance p*. Intuitively, we will write $M \sqsubseteq_p N$ if N can simulate M with a tolerance p. Thus, $\sqsubseteq_0$ will coincide with the standard weak probabilistic simulation [2,1], whereas $\sqsubseteq_1$ should be equal to $\mathtt{pTCWS} \times \mathtt{pTCWS}$.

In a probabilistic setting, the definition of weak transition is somewhat complicated by the fact that (strong) transitions take processes (in our case networks) to distributions; consequently if we are to use weak transitions $\stackrel{\alpha}{\Longrightarrow}$, which abstract away from non-observable actions, then we need to generalise transitions, so that they take (sub-)distributions to (sub-)distributions.

For a network M and a distribution Δ, we write $M \stackrel{\hat{\tau}}{\longrightarrow} \Delta$ if either $M \stackrel{\tau}{\longrightarrow} \Delta$ or $\Delta = \overline{M}$. Then, for $\alpha \neq \tau$, we write $M \stackrel{\hat{\alpha}}{\longrightarrow} \Delta$ if $M \stackrel{\alpha}{\longrightarrow} \Delta$. Relation $\stackrel{\hat{\alpha}}{\longrightarrow}$ is extended to model transitions from sub-distributions to sub-distributions. For a sub-distribution $\Delta = \sum_{i \in I} p_i \overline{M_i}$, we write $\Delta \stackrel{\hat{\alpha}}{\longrightarrow} \Theta$ if there is a set $J \subseteq I$ such that $M_j \stackrel{\hat{\alpha}}{\longrightarrow} \Theta_j$ for all $j \in J$, $M_i \not\stackrel{\hat{\alpha}}{\longrightarrow}$, for all $i \in I \setminus J$, and $\Theta = \sum_{j \in J} p_j \Theta_j$. Note that if $\alpha \neq \tau$ then this definition admits that only some networks in the support of Δ make the $\stackrel{\hat{\alpha}}{\longrightarrow}$ transition. Then, we define $\stackrel{\hat{\tau}}{\Longrightarrow} = (\stackrel{\hat{\tau}}{\longrightarrow})^*$, while for $\alpha \neq \tau$ we let $\stackrel{\hat{\alpha}}{\Longrightarrow}$ denote $\stackrel{\hat{\tau}}{\Longrightarrow}\stackrel{\hat{\alpha}}{\longrightarrow}\stackrel{\hat{\tau}}{\Longrightarrow}$.

In order to define our notion of simulation with tolerance, we adapt the concept of *weak bisimulation metric* of Desharnais et al.'s [10]. In [10], the behavioural distance between systems is measured by means of suitable *pseudometrics*, namely symmetric functions assigning a numeric value to any pair of systems. Here, we define asymmetric variants, called *pseudoquasimetrics*, measuring the tolerance of the simulation between networks. Both approaches require the lifting of these functions to distributions. In [10], this is realised by means of linear programs, relying on the symmetry of pseudometrics. Since pseudoquasimetrics are not symmetric, we need a different technique. Thus, to this end, we adopt the notions of matching [26] and Kantorovich lifting [7].

Definition 6 (Pseudoquasimetric). A function $d\colon \mathtt{pTCWS} \times \mathtt{pTCWS} \to [0,1]$ is a *1-bounded pseudoquasimetric* over $\mathtt{pTCWS}$ if (i) $d(M,M) = 0$ for all $M \in \mathtt{pTCWS}$, and (ii) $d(M,N) \leq d(M,O) + d(O,N)$ for all $M, N, O \in \mathtt{pTCWS}$.

Definition 7 (Matching). Given a pair of distributions $(\Delta, \Theta) \in \mathcal{D}(\mathtt{pTCWS}) \times \mathcal{D}(\mathtt{pTCWS})$, a *matching* of (Δ, Θ) is a distribution $\omega \in \mathcal{D}(\mathtt{pTCWS} \times \mathtt{pTCWS})$ s.t.: (i) $\sum_{N \in \mathtt{pTCWS}} \omega(M, N) = \Delta(M)$, for all $M \in \mathtt{pTCWS}$, and (ii) $\sum_{M \in \mathtt{pTCWS}} \omega(M, N) = \Theta(N)$, for all $N \in \mathtt{pTCWS}$. $\Omega(\Delta, \Theta)$ denotes the set of all matchings for (Δ, Θ).

A matching for (Δ, Θ) may be understood as a transportation schedule for the shipment of probability mass from Δ to Θ [26].

Definition 8 (Kantorovich lifting). Let $d\colon \mathtt{pTCWS} \times \mathtt{pTCWS} \to [0,1]$ be a pseudoquasimetric. The *Kantorovich lifting* of d is the function $\mathbf{K}(d)\colon \mathcal{D}(\mathtt{pTCWS}) \times \mathcal{D}(\mathtt{pTCWS}) \to [0,1]$ defined as:

$$\mathbf{K}(d)(\Delta, \Theta) \stackrel{\text{def}}{=} \min_{\omega \in \Omega(\Delta,\Theta)} \sum_{M,N \in \mathtt{pTCWS}} \omega(M, N) \cdot d(M, N) \ .$$

Note that since we are considering only distributions with finite support, the minimum over the set of matchings $\Omega(\Delta, \Theta)$ is well defined.

Definition 9 (Weak simulation quasimetric). We say that a pseudoquasimetric $d\colon \mathtt{pTCWS} \times \mathtt{pTCWS} \to [0,1]$ is a *weak simulation quasimetric* if for all networks $M, N \in \mathtt{pTCWS}$, with $d(M, N) < 1$, whenever $M \xrightarrow{\alpha} \Delta$ there is a subdistribution Θ such that $N \stackrel{\hat{\alpha}}{\Longrightarrow} \Theta$ and $\mathbf{K}(d)(\Delta\,, \Theta + (1- |\Theta|)\mathsf{Dead}) \leq d(M, N)$.

In the previous definition, if $|\Theta| < 1$ then, with probability $1- |\Theta|$, there is no way to simulate the behaviour of any network in the support of Δ (the special network Dead does not perform any action).

As expected, the kernel of a weak simulation quasimetric is a weak probabilistic simulation [2,1].

Proposition 1. Let d be a weak simulation quasimetric. The binary relation $\{(M, N) : d(M, N) = 0\} \subseteq \mathtt{pTCWS} \times \mathtt{pTCWS}$ is a weak probabilistic simulation.

A crucial result in our construction process is the existence of the minimal weak simulation quasimetric, which can be viewed as the asymmetric counterpart of the minimal weak bisimulation metric [10].

Theorem 1. There is a weak simulation quasimetric $\mathbf{d}$ s.t. $\mathbf{d}(M, N) \leq d(M, N)$ for all weak simulation quasimetrics d and all networks $M, N \in \mathtt{pTCWS}$.

Now, we have all ingredients to define our simulation with tolerance p.

Definition 10 (Weak simulation with tolerance). Let $p \in [0,1]$, we say that N *simulates* M *with tolerance* p, written $M \sqsubseteq_p N$, iff $\mathbf{d}(M, N) = q$, for some $q \leq p$. We write $M \simeq_p N$ if both $M \sqsubseteq_p N$ and $N \sqsubseteq_p M$.

Since the minimum weak simulation quasimetric $\mathbf{d}$ satisfies the triangle inequality, our simulation relation is trivially transitive in an additive sense:

Proposition 2 (Transitivity). $M \sqsubseteq_p N$ and $N \sqsubseteq_q O$ imply $M \sqsubseteq_{p+q} O$.

As expected, if $M \stackrel{\hat{\tau}}{\Longrightarrow} \Delta$ then M can simulate all networks in $\lceil \Delta \rceil$.

Proposition 3. If $M \stackrel{\hat{\tau}}{\Longrightarrow} (1-q)\overline{N} + q\Delta$, for some $\Delta \in \mathcal{D}(\mathtt{pTCWS})$, then $N \sqsubseteq_q M$.

Clearly the transitivity property is quite useful when doing algebraic reasoning. However, we can derive a better tolerance when concatenating two simulations, if one of them is derived by an application of Proposition 3.

Proposition 4. If $M \sqsubseteq_p N$ and $O \stackrel{\hat{\tau}}{\Longrightarrow} (1-q)\overline{N} + q\Delta$, for some $\Delta \in \mathcal{D}(\mathtt{pTCWS})$, then $M \sqsubseteq_{p(1-q)+q} O$.

Intuitively, in the simulation between M and N the tolerance p must be weighted by taking into consideration that O may evolve into N with a probability $(1-q)$.

In order to understand the intuition behind our weak simulation with tolerance, we report here a few simple *algebraic laws* (we recall that $1{:}P = P$).

Proposition 5 (Simple algebraic laws).
1. $n[P]^\mu \ \sqsubseteq_{1-p} \ n[\tau.(P \oplus_p Q)]^\mu$
2. $n[Q]^\mu \ \sqsubseteq_r \ \ n[\tau.(\tau.(P \oplus_q Q) \oplus_p R)]^\mu$, with $r = (1-p) + pq$
3. $n[!\langle v\rangle.\big(\tau.(P \oplus_q \tau.P) \oplus_p Q\big)]^\mu \ \simeq_0 \ n[!\langle v\rangle.(P \ \oplus_p \ \tau.Q)]^\mu$
4. $n[!\langle v\rangle.!\langle w\rangle]^\mu \ \sqsubseteq_r \ n[\tau.\big(!\langle v\rangle.\tau.(!\langle w\rangle \oplus_q P) \ \oplus_p \ Q\big)]^\mu$, with $r = 1 - pq$.

The first law is straightforward. The second law is a generalisation of the first one where the right-hand side must resolve two probabilistic choices in order to simulate the left-hand side. The third law is an adaptation of the CCS tau-law $\tau.P = P$ in a distributed and probabilistic setting. Similarly, the fourth law reminds a probabilistic and distributed variant of the tau-law $a.(\tau.(P + \tau.Q)) + a.Q = a.(P + \tau.Q)$. This law gives an example of a probabilistic simulation involving sequences of actions.

A crucial property of our simulation is the possibility to reason on parallel networks in a *compositional* manner. Thus, if $M_1 \sqsubseteq_{p_1} N_1$ and $M_2 \sqsubseteq_{p_2} N_2$ then $M_1 \mid M_2 \sqsubseteq_p N_1 \mid N_2$ for some p depending on p_1 and p_2; the intuition being that if one fixes the maximal tolerance p between $M_1 \mid M_2$ and $N_1 \mid N_2$, then there are tolerances p_i between M_i and N_i ensuring that the tolerance p is respected. Following this intuition, several compositional criteria for bisimulation metrics can be found in the literature [10,12,13,14,15]. Here, we show that our weak simulation with tolerance complies with *non-expansiveness*: one of the strongest criteria, requiring $p \leq p_1 + p_2$.

Theorem 2 (Non-expansiveness law). $M_1 \sqsubseteq_{p_1} N_1$ and $M_2 \sqsubseteq_{p_2} N_2$ entails $M_1 \mid M_2 \sqsubseteq_{p_1+p_2} N_1 \mid N_2$.

Another useful property is that a network is simulated by a probabilistic choice whenever it is simulated by all components.

Proposition 6 (Additive law). Let $M \sqsubseteq_{s_i} n[P_i]^\mu \mid N$, for all $i \in I$, with I a finite index set. Then, $M \sqsubseteq_r n[\tau.\bigoplus_{i\in I} p_i{:}P_i]^\mu \mid N$, for $r = \sum_{i\in I} p_i s_i$.

Finally, we report a number of algebraic laws that will be useful in the next section when analysing gossip protocols.

Proposition 7 (Further algebraic laws).
1. $n[\sigma^k.\mathsf{nil}]^\mu \ \simeq_0 \ n[\mathsf{nil}]^\mu$

2. $\prod_{i\in I} m_i[P_i]^{\mu_i} \simeq_r \prod_{j\in J} n_j[Q_j]^{\nu_j}$ entails $\prod_{i\in I} m_i[\sigma.P_i]^{\mu_i} \simeq_r \prod_{j\in J} n_j[\sigma.Q_j]^{\nu_j}$
3. $n[?(x).C]^{\mu} \simeq_0 n[\sigma.?(x).C]^{\mu}$, if nodes in μ do not send in the current round
4. $m[\mathsf{nil}]^{\mu} \mid \prod_{i\in I} n_i[P_i]^{\mu_i} \sqsubseteq_0 m[\tau.(!\langle v\rangle \oplus_p \mathsf{nil})]^{\mu} \mid \prod_{i\in I} n_i[P_i]^{\mu_i}$ if $\mu \subseteq \bigcup_{i\in I} n_i$, and for all $n_i \in \mu$ it holds that $P_i \neq \lfloor ?(x).C \rfloor D$.

Intuitively: (1) nil does not prevent time passing; (2) equalities are preserved underneath σ prefixes; (3) receptions will timeout if there are not senders around; (4) broascast has no effect if there are not receivers around.

4 A case study: reasoning on gossip protocols

The baseline model for our case study is gossiping without communication collisions, where all nodes are perfectly synchronised. For the sake of clarity, communication proceeds in synchronous rounds: a node can transmit or receive only one message per round. In our implementation, rounds are separated by σ-actions.

The processes involved in the protocol are the following:

$$\mathsf{snd}\langle u\rangle_{p_\mathrm{g}} \stackrel{\text{def}}{=} \tau.(!\langle u\rangle \oplus_{p_\mathrm{g}} \mathsf{nil}) \quad \mathsf{fwd}_{p_\mathrm{g}} \stackrel{\text{def}}{=} ?(x).\mathsf{resnd}\langle x\rangle_{p_\mathrm{g}} \quad \mathsf{resnd}\langle u\rangle_{p_\mathrm{g}} \stackrel{\text{def}}{=} \sigma.\mathsf{snd}\langle u\rangle_{p_\mathrm{g}}.$$

A sender broadcasts with a gossip probability p_g, whereas a forwarder rebroadcasts the received value, in the subsequent round, with the same probability.

We apply our simulation theory to develop algebraic reasonings on *message propagation*. As an introductory example, let us consider a fragment of a network with two sender nodes, m_1 and m_2, and two forwarder nodes, n_1 and n_2 which are both neighbours of m_1 and m_2. Then, the following holds:

$$\begin{array}{c} m_1[\mathsf{snd}\langle u\rangle_{p_1}]^{\nu} \mid m_2[\mathsf{snd}\langle u\rangle_{p_2}]^{\nu} \mid n_1[\mathsf{fwd}_q]^{\nu_1} \mid n_2[\mathsf{fwd}_r]^{\nu_2} \quad {}_s\!\sqsupseteq \\ m_1[\mathsf{nil}]^{\nu} \mid m_2[\mathsf{nil}]^{\nu} \mid n_1[\mathsf{resnd}\langle u\rangle_q]^{\nu_1} \mid n_2[\mathsf{resnd}\langle u\rangle_r]^{\nu_2} \end{array}$$

with tolerance $s = (1-p_1)(1-p_2)$. Here, the network on the left-hand-side evolves by performing two τ-actions (via rule (ShhSnd)). Thus, the algebraic law follows by an application of Proposition 3 being $1-s$ the probability that the message u is broadcast to both forwarders.

This simple law can be generalised to an arbitrary number of senders and forwarders, under the hypothesis that parallel contexts are unable to receive messages in the current round. The following theorem relies on Proposition 3.

Theorem 3 (Message propagation). Let I and J be pairwise disjoint subsets of $\mathbb{N}$. Let M be a well-formed network defined as

$$M \equiv N \mid \prod_{i\in I} m_i[\mathsf{snd}\langle v\rangle_{p_i}]^{\nu_{m_i}} \mid \prod_{j\in J} n_j[\mathsf{fwd}_{q_j}]^{\nu_{n_j}}$$

such that, for all $i \in I$:
- $\{n_j : j \in J\} \subseteq \nu_{m_i} \subseteq \mathsf{nds}(M)$, and
- the nodes in $\nu_{m_i} \cap \mathsf{nds}(N)$ cannot receive in the current round.

Then, $M \;{}_r\!\sqsupseteq\; N \mid \prod_{i\in I} m_i[\mathsf{nil}]^{\nu_{m_i}} \mid \prod_{j\in J} n_j[\mathsf{resnd}\langle v\rangle_{q_j}]^{\nu_{n_j}}$, with $r = \prod_{i\in I}(1-p_i)$.

Theorem 3 represents an effective tool to deal with message propagation in gossip networks. However, it requires that all forwarders n_j should be in the neighbouring of all senders m_i (constraint $\{n_j : j \in J\} \subseteq \nu_{m_i}$), which may represent a limitation in many cases. Consider, for example, a simple gossiping network GSP, with gossip probability p, composed by two source nodes s_1 and s_2, a destination node d, and three intermediate nodes n_1, n_2 and n_3:

$$\text{GSP} \stackrel{\text{def}}{=} \prod_{i=1}^{2} s_i[\mathsf{snd}\langle v\rangle_p]^{\nu_{s_i}} \mid \prod_{i=1}^{3} n_i[\mathsf{fwd}_p]^{\nu_{n_i}} \mid d[\mathsf{fwd}_1]^{\nu_d} \tag{1}$$

with topology $\nu_{s_1} = \{n_1\}$, $\nu_{s_2} = \{n_1, n_2\}$, ν_{n_1}=$\{s_1, s_2, n_3\}$, ν_{n_2}=$\{s_2, n_3\}$, $\nu_{n_3} = \{n_1, n_2, d\}$ and $n_3 \in \nu_d$.

Here, we would like to estimate the distance between GSP, and a network DONE, in which the message v has been delivered to the destination node d.

$$\text{DONE} \stackrel{\text{def}}{=} \prod_{i=1}^{2} s_i[\mathsf{nil}]^{\nu_{s_i}} \mid \prod_{i=1}^{3} n_i[\mathsf{nil}]^{\nu_{n_i}} \mid d[\sigma^3.\,\mathsf{snd}\langle v\rangle_1]^{\nu_d} \tag{2}$$

Unfortunately, we cannot directly apply Theorem 3 to capture this message propagation because node s_2, unlike s_1, can transmit to both n_1 and n_2. In this case, before applying Theorem 3, we would need a result to *compose estimates* of partial networks. More precisely, a result which would allow us to take into account, in the calculation of the tolerance, both the probability that a sender transmits and the probability that the same sender does not transmit. The following result follows from Proposition 6.

Theorem 4 (Composing networks). If $M \stackrel{\sigma}{\not\rightarrow}$ then

$$N \mid m[\mathsf{snd}\langle v\rangle_p]^{\nu_m} \mid \prod_{j\in J} n_j[\lfloor ?(x_j).P_j \rfloor Q_j]^{\nu_{n_j}} \;{}_r\sqsupseteq\; M$$

with tolerance $r = ps_1 + (1-p)s_2$, whenever
- $N \mid m[\mathsf{nil}]^{\nu_m} \mid \prod_{j\in J} n_j[\{^v/_{x_j}\}P_j]^{\nu_{n_j}} \;{}_{s_1}\sqsupseteq\; M$
- $N \mid m[\mathsf{nil}]^{\nu_m} \mid \prod_{j\in J} n_j[\lfloor ?(x_j).P_j \rfloor Q_j]^{\nu_{n_j}} \;{}_{s_2}\sqsupseteq\; M$
- $\{n_j : j \in J\} \subseteq \nu_m \subseteq \{n_j : j \in J\} \cup \mathsf{nds}(N)$
- nodes in $\nu_m \cap \mathsf{nds}(N)$ cannot receive in the current round.[3]

Intuitively: (i) in the network $N \mid m[\mathsf{snd}\langle v\rangle_p]^{\nu_m} \mid \prod_{j\in J} n_j[\lfloor ?(x_j).P_j \rfloor Q_j]^{\nu_{n_j}}$ node m has not performed yet the τ-action that resolves the probabilistic choice between broadcasting v or not; (ii) in $N \mid m[\mathsf{nil}]^{\nu_m} \mid \prod_{j\in J} n_j[\{^v/_{x_j}\}P_j]^{\nu_{n_j}}$ node m has resolved the probabilistic choice deciding to broadcast v; (iii) finally, in the network $N \mid m[\mathsf{nil}]^{\nu_m} \mid \prod_{j\in J} n_j[\lfloor ?(x_j).P_j \rfloor Q_j]^{\nu_{n_j}}$ node m has has resolved the probabilistic choice deciding not to broadcast v.

Now, we have all algebraic tools to compute an estimation of the tolerance r, such that GSP ${}_r\sqsupseteq$ DONE. In practise, we will compute the tolerance for two partial networks and then will use Theorem 4 to compose the two tolerances.

For verification reasons we assume that the environment contains a node *test*, close to the destination node, i.e. $\nu_d = \{n_3, test\}$, to test successful gossiping. For simplicity, the *test* node can receive messages but it cannot transmit.

[3] We could generalise the result to take into account more senders at the same time. This would not add expressiveness, it would just speed up the reduction process.

As a first step, we compute an estimation for the network GSP in which the sender s_2 *has already broadcast* the message v to its neighbours n_1 and n_2. To this end, we derive the following chain of similarities by applying, in sequence, (i) Proposition 7(4), (ii) Proposition 7(3), (iii) Theorem 3 and Proposition 7(2), (iv) Proposition 7(1) and Proposition 7(3), (v) Theorem 3 and Proposition 7(2), and (vi) Proposition 7(1). In all steps, we have reasoned in a compositional manner, up to common parallel components (Theorem 2).

$$
\begin{aligned}
& s_1[\mathsf{snd}\langle v\rangle_p]^{\nu_{s_1}} \mid s_2[\mathsf{nil}]^{\nu_{s_2}} \mid \textstyle\prod_{i=1}^{2} n_i[\mathsf{resnd}\langle v\rangle_p]^{\nu_{n_i}} \mid n_3[\mathsf{fwd}_p]^{\nu_{n_3}} \mid d[\mathsf{fwd}_1]^{\nu_d} \\
\sqsupseteq_{0} & \textstyle\prod_{i=1}^{2} s_i[\mathsf{nil}]^{\nu_{s_i}} \mid \prod_{i=1}^{2} n_i[\mathsf{resnd}\langle v\rangle_p]^{\nu_{n_i}} \mid n_3[\mathsf{fwd}_p]^{\nu_{n_3}} \mid d[\mathsf{fwd}_1]^{\nu_d} \\
\sqsupseteq_{0} & \textstyle\prod_{i=1}^{2} s_i[\mathsf{nil}]^{\nu_{s_i}} \mid \prod_{i=1}^{2} n_i[\sigma.\mathsf{snd}\langle v\rangle_p]^{\nu_{n_i}} \mid n_3[\sigma.\mathsf{fwd}_p]^{\nu_{n_3}} \mid d[\sigma.\mathsf{fwd}_1]^{\nu_d} \\
\sqsupseteq_{(1-p)^2} & \textstyle\prod_{i=1}^{2} s_i[\mathsf{nil}]^{\nu_{s_i}} \mid \prod_{i=1}^{2} n_i[\sigma.\mathsf{nil}]^{\nu_{n_i}} \mid n_3[\sigma.\mathsf{resnd}\langle v\rangle_p]^{\nu_{n_3}} \mid d[\sigma.\mathsf{fwd}_1]^{\nu_d} \\
\sqsupseteq_{0} & \textstyle\prod_{i=1}^{2} s_i[\mathsf{nil}]^{\nu_{s_i}} \mid \prod_{i=1}^{2} n_i[\mathsf{nil}]^{\nu_{n_i}} \mid n_3[\sigma^2.\,\mathsf{snd}\langle v\rangle_p]^{\nu_{n_3}} \mid d[\sigma^2.\,\mathsf{fwd}_1]^{\nu_d} \\
\sqsupseteq_{1-p} & \textstyle\prod_{i=1}^{2} s_i[\mathsf{nil}]^{\nu_{s_i}} \mid \prod_{i=1}^{2} n_i[\mathsf{nil}]^{\nu_{n_i}} \mid n_3[\sigma^2.\,\mathsf{nil}]^{\nu_{n_3}} \mid d[\sigma^2.\,\mathsf{resnd}\langle v\rangle_1]^{\nu_d} \\
\sqsupseteq_{0} & \textstyle\prod_{i=1}^{2} s_i[\mathsf{nil}]^{\nu_{s_i}} \mid \prod_{i=1}^{3} n_i[\mathsf{nil}]^{\nu_{n_i}} \mid d[\sigma^3.\,\mathsf{snd}\langle v\rangle_1]^{\nu_d} \\
= & \ \text{DONE} \ .
\end{aligned}
$$

Then, by more applications of Proposition 2 and Proposition 7(1), one application of Proposition 4, and one application of Proposition 7(2) we derive:

$$
\begin{aligned}
& s_1[\mathsf{snd}\langle v\rangle_p]^{\nu_{s_1}} \mid s_2[\mathsf{nil}]^{\nu_{s_2}} \mid \textstyle\prod_{i=1}^{2} n_i[\mathsf{resnd}\langle v\rangle_p]^{\nu_{n_i}} \mid n_3[\mathsf{fwd}_p]^{\nu_{n_3}} \mid d[\mathsf{fwd}_1]^{\nu_d} \\
& \qquad \sqsupseteq_{1-2p^2+p^3} \ \text{DONE}
\end{aligned}
\tag{3}
$$

with tolerance $1-2p^2+p^3$, obtained by solving the expression $(1-p)(1-(1-p)^2)+(1-p)^2$.

Similarly, we compute an estimation of the tolerance which allows the network GSP, in which the sender s_2 *did not broadcast* the message v to its neighbours, to simulate the network DONE. To this end, we derive the following chain of similarities by applying, in sequence, (i) Theorem 3 and Proposition 7(3), (ii) Proposition 7(3), (iii) Theorem 3 and Proposition 7(2), (iv) Proposition 7(1) and Proposition 7(3), (v) Theorem 3 and Proposition 7(2), and (vi) Proposition 7(1) and Proposition 7(4). In all steps, we have reasoned up to common parallel components (Theorem 2).

$$
\begin{aligned}
& s_1[\mathsf{snd}\langle v\rangle_p]^{\nu_{s_1}} \mid s_2[\mathsf{nil}]^{\nu_{s_2}} \mid \textstyle\prod_{i=1}^{3} n_i[\mathsf{fwd}_p]^{\nu_{n_i}} \mid d[\mathsf{fwd}_1]^{\nu_d} \\
\sqsupseteq_{1-p} & \textstyle\prod_{i=1}^{2} s_i[\mathsf{nil}]^{\nu_{s_i}} \mid n_1[\mathsf{resnd}\langle v\rangle_p]^{\nu_{n_1}} \mid \prod_{i=2}^{3} n_i[\mathsf{fwd}_p]^{\nu_{n_i}} \mid d[\mathsf{fwd}_1]^{\nu_d} \\
\sqsupseteq_{0} & \textstyle\prod_{i=1}^{2} s_i[\mathsf{nil}]^{\nu_{s_i}} \mid n_1[\sigma.\mathsf{snd}\langle v\rangle_p]^{\nu_{n_1}} \mid \prod_{i=2}^{3} n_i[\sigma.\mathsf{fwd}_p]^{\nu_{n_i}} \mid d[\sigma.\mathsf{fwd}_1]^{\nu_d} \\
\sqsupseteq_{1-p} & \textstyle\prod_{i=1}^{2} s_i[\mathsf{nil}]^{\nu_{s_i}} \mid n_1[\sigma.\mathsf{nil}]^{\nu_{n_1}} \mid n_2[\sigma.\mathsf{fwd}_p]^{\nu_{n_2}} \mid n_3[\sigma.\mathsf{resnd}\langle v\rangle_p]^{\nu_{n_3}} \mid d[\sigma.\mathsf{fwd}_1]^{\nu_d} \\
\sqsupseteq_{0} & \textstyle\prod_{i=1}^{2} s_i[\mathsf{nil}]^{\nu_{s_i}} \mid n_1[\mathsf{nil}]^{\nu_{n_1}} \mid n_2[\sigma^2.\,\mathsf{fwd}_p]^{\nu_{n_2}} \mid n_3[\sigma^2.\,\mathsf{snd}\langle v\rangle_p]^{\nu_{n_3}} \mid d[\sigma^2.\,\mathsf{fwd}_1]^{\nu_d} \\
\sqsupseteq_{1-p} & \textstyle\prod_{i=1}^{2} s_i[\mathsf{nil}]^{\nu_{s_i}} \mid n_1[\mathsf{nil}]^{\nu_{n_1}} \mid n_2[\sigma^2.\,\mathsf{resnd}\langle v\rangle_p]^{\nu_{n_2}} \mid n_3[\sigma^2.]^{\nu_{n_3}} \mid d[\sigma^2.\,\mathsf{resnd}\langle v\rangle_1]^{\nu_d} \\
\sqsupseteq_{0} & \textstyle\prod_{i=1}^{2} s_i[\mathsf{nil}]^{\nu_{s_i}} \mid \prod_{i=1}^{3} n_i[\mathsf{nil}]^{\nu_{n_i}} \mid d[\sigma^3.\mathsf{snd}\langle v\rangle_1]^{\nu_d} \\
= & \ \text{DONE} \ .
\end{aligned}
$$

Then, by more applications of Proposition 2 and Proposition 7(1), one application of Proposition 4, and one application of Proposition 7(2) we derive:

$$s_1[\mathsf{snd}\langle v\rangle_p]^{\nu_{s_1}} \mid s_2[\mathsf{nil}]^{\nu_{s_2}} \mid \prod_{i=1}^{3} n_i[\mathsf{fwd}_p]^{\nu_{n_i}} \mid d[\mathsf{fwd}_1]^{\nu_d} \;\; {}_{1-p^3}\sqsupseteq \;\; \text{DONE} \;. \quad (4)$$

Finally, we can apply Theorem 4 to (3) and (4) to derive the following estimation for the tolerance:

$$\text{GSP} \;\; {}_{1-(3p^3-2p^4)}\sqsupseteq \;\; \text{DONE}$$

Since the tolerance is $1-(3p^3-2p^4)$, it follows that the gossip network GSP will succeed in propagating the messages to the destination d, with probability at least $3p^3-2p^4$. Thus, for instance, for a gossip probability $p = 0.8$ the destination will receive the message with probability 0.716, with a margin of 10%. For $p = 0.85$ the probability at the destination increases to 0.798, with a margin of 6%; while for $p = 0.9$ the probability at destination rises to 0.88, with a difference of only 2%. So, $p = 0.9$ can be considered the threshold of our small network.[4]

Note that in the previous example both messages may reach the destination node in exactly three rounds. However, more generally, we could have different message propagation paths in the same network which might take a different amount of time to be traversed. The algebraic tools we developed up to now do not allow us to deal with paths of different lengths.

As an example, we would like to estimate the distance between the network

$$\text{GSP}_2 \stackrel{\text{def}}{=} s_1[\mathsf{snd}\langle v\rangle_1]^{\nu_{s_1}} \mid s_2[\mathsf{snd}\langle v\rangle_p]^{\nu_{s_2}} \mid n[\mathsf{fwd}_p]^{\nu_n} \mid d[\mathsf{fwd}_1]^{\nu_d}$$

with topology $\nu_{s_1} = \{d\}$, $\nu_{s_2} = \{n\}$, ν_n=$\{s_2, d\}$ and $\nu_d = \{s_1, n, test\}$, and the networks defined as follows:

$$\text{DONE}_2 \stackrel{\text{def}}{=} s_1[\mathsf{nil}]^{\nu_{s_1}} \mid s_2[\mathsf{nil}]^{\nu_{s_2}} \mid n[\mathsf{nil}]^{\nu_n} \mid d[\tau.(\sigma.\mathsf{snd}\langle v\rangle_1 \oplus_p \sigma^2.\mathsf{snd}\langle v\rangle_1)]^{\nu_d}$$

in which the message v propagated up to the destination node d following two different paths. Thus, d will probabilistically choose between broadcasting v after one or two rounds.

The following result provide the missing instrument.

Theorem 5 (Composing paths). Let M be a well-formed network. Then,

$$M \mid m[\tau.\bigoplus_{i\in I} p_i{:}Q_i]^{\nu_m} \;\; {}_r\sqsupseteq \;\; \prod_{j\in J} n_j[\mathsf{nil}]^{\nu_{n_j}} \mid d[\tau.\bigoplus_{i\in I} p_i{:}P_i]^{\nu_d}$$

with $r = \sum_{i\in I} p_i s_i$, whenever:

- $M \mid m[Q_i]^{\nu_m} \;\; {}_{s_i}\sqsupseteq \;\; \prod_{j\in J} n_j[\mathsf{nil}]^{\nu_{n_j}} \mid d[P_i]^{\nu_d}$, for any $i \in I$;
- $\nu_m \subseteq \mathsf{nds}(M)$.

As a first step, we compute an estimation of the tolerance which allows GSP_2 to simulate the first probabilistic behaviour of DONE_2. To this end, we derive the following chain of similarities by applying, in sequence, (i) Theorem 3, (ii)

[4] Had we had more senders we would have estimated a better threshold.

again Theorem 3, (iii) Proposition 7(2) and Proposition 7(4). In all steps, we reason up to parallel components (Theorem 2).

$$\begin{aligned} & s_1[\mathsf{snd}\langle v\rangle_1]^{\nu_{s_1}} \mid s_2[\mathsf{snd}\langle v\rangle_p]^{\nu_{s_2}} \mid n[\mathsf{fwd}_p]^{\nu_n} \mid d[\mathsf{fwd}_1]^{\nu_d} \\ {}_0\sqsupseteq\ & s_1[\mathsf{nil}]^{\nu_{s_1}} \mid s_2[\mathsf{snd}\langle v\rangle_1]^{\nu_{s_2}} \mid n[\mathsf{fwd}_p]^{\nu_n} \mid d[\mathsf{resnd}\langle v\rangle_1]^{\nu_d} \\ {}_{1-p}\sqsupseteq\ & s_1[\mathsf{nil}]^{\nu_{s_1}} \mid s_2[\mathsf{nil}]^{\nu_{s_2}} \mid n[\mathsf{resnd}\langle v\rangle_p]^{\nu_n} \mid d[\mathsf{resnd}\langle v\rangle_1]^{\nu_d} \\ {}_0\sqsupseteq\ & s_1[\mathsf{nil}]^{\nu_{s_1}} \mid s_2[\mathsf{nil}]^{\nu_{s_2}} \mid n[\mathsf{nil}]^{\nu_n} \mid d[\sigma.\mathsf{snd}\langle v\rangle_1]^{\nu_d}. \end{aligned}$$

By an application of Proposition 2 we derive:

$$\begin{aligned} & s_1[\mathsf{snd}\langle v\rangle_1]^{\nu_{s_1}} \mid s_2[\mathsf{snd}\langle v\rangle_p]^{\nu_{s_2}} \mid n[\mathsf{fwd}_p]^{\nu_n} \mid d[\mathsf{fwd}_1]^{\nu_d} \\ {}_{1-p}\sqsupseteq\ & s_1[\mathsf{nil}]^{\nu_{s_1}} \mid s_2[\mathsf{nil}]^{\nu_{s_2}} \mid n[\mathsf{nil}]^{\nu_n} \mid d[\sigma.\mathsf{snd}\langle v\rangle_1]^{\nu_d}. \end{aligned} \tag{5}$$

Then, we compute an estimation of the tolerance which allows the network GSP_2 to simulate the second probabilistic behaviour of DONE_2. To this end, we derive the following chain of similarities by applying, in sequence, (i) Theorem 3, (ii) Theorem 3 again, Proposition 7(1), Proposition 7(2) and Proposition 7(3). Again, in all steps, we have reasoned up to parallel components (Theorem 2).

$$\begin{aligned} & s_1[\mathsf{nil}]^{\nu_{s_1}} \mid s_2[\mathsf{snd}\langle v\rangle_p]^{\nu_{s_2}} \mid n[\mathsf{fwd}_p]^{\nu_n} \mid d[\mathsf{fwd}_1]^{\nu_d} \\ {}_{1-p}\sqsupseteq\ & s_1[\mathsf{nil}]^{\nu_{s_1}} \mid s_2[\mathsf{nil}]^{\nu_{s_2}} \mid n[\mathsf{resnd}\langle v\rangle_p]^{\nu_n} \mid d[\mathsf{fwd}_1]^{\nu_d} \\ {}_{1-p}\sqsupseteq\ & s_1[\mathsf{nil}]^{\nu_{s_1}} \mid s_2[\mathsf{nil}]^{\nu_{s_2}} \mid n[\mathsf{nil}]^{\nu_n} \mid d[\sigma^2.\mathsf{snd}\langle v\rangle_1]^{\nu_d} \ . \end{aligned}$$

Then, by more applications of Proposition 2 and one application of Proposition 4 we derive:

$$\begin{aligned} & s_1[\mathsf{nil}]^{\nu_{s_1}} \mid s_2[\mathsf{snd}\langle v\rangle_p]^{\nu_{s_2}} \mid n[\mathsf{fwd}_p]^{\nu_n} \mid d[\mathsf{fwd}_1]^{\nu_d} \\ {}_{1-p^2}\sqsupseteq\ & s_1[\mathsf{nil}]^{\nu_{s_1}} \mid s_2[\mathsf{nil}]^{\nu_{s_2}} \mid n[\mathsf{nil}]^{\nu_n} \mid d[\sigma^2.\mathsf{snd}\langle v\rangle_1]^{\nu_d} \ . \end{aligned} \tag{6}$$

Finally, we can apply Theorem 5 to (5) and (6) to derive

$$\mathrm{GSP}_2 \ \ {}_r\sqsupseteq \ \ \mathrm{DONE}_2$$

with $r = p(1-p) + (1-p)(1-p^2)$. Thus, the network GSP_2 will succeed in transmitting both messages v to the destination d, with probability at least $1-r$.

We conclude by observing that, in order to deal with paths of different length, one should apply Theorem 5 for all possible paths.

5 Conclusions, related and future work

We have introduced the notion of *weak simulation quasimetric* as a means to define *weak simulation with tolerance*, i.e. a *compositional* simulation theory to express that a probabilistic system may be simulated by another one with a given tolerance measuring the distance between the two systems. Basically, weak simulation quasimetric is the asymmetric counterpart of *weak bisimulation metric* [10], and the quantitative analogous of *weak simulation preorder* [2,1].

We applied our proposal to develop an *algebraic theory* to estimate the performance of gossip networks in terms of the probability to successfully propagate messages up to the desired destination. The algebraic theory is compositional as it allows us to estimate the performance of gossip networks in terms of the behavioural distance of its sub-networks.

Our work has been inspired by [9,10,4,6], where the notion of behavioural distance between two probabilistic systems is formalised in terms of the notion of bisimulation metric. Bisimulation metric works fine for systems being approximately equivalent. However, when the simulation game works only in one direction, as in the gossip protocols analysed in the current paper, an asymmetric notion of simulation pseudometric is required.

The current paper is the ideal continuation of [19]. In that paper, the authors developed a notion of *simulation up to probability* to measure the closeness rather than the distance between two probabilistic systems. Then, as in here, simulation up to probability has been used to provide an algebraic theory to evaluate the performance of gossip networks. Despite the similarity of the two simulation theories, the simulation up to probability has a number of limitations that have motivated the current work: (i) the simulation up to probability is not transitive, while simulation quasimetrics are transitive by definition; (ii) in order to work with a transitive relation, paper [19] introduces an auxiliary rooted simulation which is much stronger than the main definition; (iii) that rooted simulation (and hence the simulation up to probability) is not suitable to compose estimates originating from paths with different lengths (as we do here by means of Theorem 5), and, more generally, to deal with more transmissions.

A nice survey of formal verification techniques for the analysis of gossip protocols appears in [3]. Probabilistic model-checking has been used in [11] to study the influence of different modelling choices on message propagation in flooding and gossip protocols, and in [18] to investigate the expected rounds of gossiping required to form a connected network and how the expected path length between nodes evolves over the execution of the protocol.

As future work, we intend to study gossip protocols with communication collisions, random delays and lossy channels. We then plan to apply our metric-based simulation theory to investigate the behaviour of IoT systems and cyber-physical systems [20,21]. In the context of probabilistic process calculi, we want to investigate which of the compositionality properties proposed in [13] hold for the operators that are usually offered by probabilistic process calculi.

Acknowledgements. We thank the anonymous reviewers for valuable comments.

References

1. Baier, C., Hermanns, H., Katoen, J.P.: Probabilistic weak simulation is decidable in polynomial time. Information Processing Letters 89(3), 123–130 (2004)
2. Baier, C., Katoen, J.P., Hermanns, H., Haverkort, B.R.: Simulation for Continuous-Time Markov Chains. In: Brim, L., Jancar, P., Kretínský, M., Kucera, A. (eds.) CONCUR 2002. LNCS, vol. 2421, pp. 338–354. Springer (2002)
3. Bakhshi, R., Bonnet, F., Fokkink, W., Haverkort, B.: Formal analysis techniques for gossiping protocols. Operating Systems Review 41(5), 28–36 (2007)

4. van Breugel, F., Worrell, J.: A behavioural pseudometric for probabilistic transition systems. Theoretical Computer Science 331(1), 115–142 (2005)
5. Cerone, A., Hennessy, M., Merro, M.: Modelling mac-layer communications in wireless systems. Logical Methods in Computer Science 11(1:18) (2015)
6. Deng, Y., Chothia, T., Palamidessi, C., Pang, J.: Metrics for Action-labelled Quantitative Transition Systems. ENTCS 153(2), 79–96 (2006)
7. Deng, Y., Du, W.: The Kantorovich Metric in Computer Science: A Brief Survey. ENTCS 253(3), 73–82 (2009)
8. Deng, Y., van Glabbeek, R.J., Hennessy, M., Morgan, C.: Characterising testing preorders for finite probabilistic processes. Logical Meth. Comput. Sci. 4(4) (2008)
9. Desharnais, J., Gupta, J., Jagadeesan, R., Panangaden, P.: Metrics for Labelled Markov Processes. Theoretical Computer Science 318(3), 323–354 (2004)
10. Desharnais, J., Jagadeesan, R., Gupta, V., Panangaden, P.: The metric analogue of weak bisimulation for probabilistic processes. In: LICS 2002. pp. 413–422 (2002)
11. Fehnker, A., Gao, P.: Formal Verification and Simulation for Performance Analysis for Probabilistic Broadcast Protocols. In: Kunz, T., Ravi, S.S. (eds.) ADHOC-NOW 2006. LNCS, vol. 4104, pp. 128–141. Springer (2006)
12. Gebler, D., Larsen, K.G., Tini, S.: Compositional Metric Reasoning with Probabilistic Process calculi. In: Pitts, A.M. (ed.) FoSSaCS 2015. LNCS, vol. 9034, pp. 230–245. Springer (2015)
13. Gebler, D., Larsen, K.G., Tini, S.: Compositional Bisimulation Metric Reasoning with Probabilistic Process Calculi. Logical Meth. Comput. Sci. 12(4) (2016)
14. Gebler, D., Tini, S.: Fixed-point Characterization of Compositionality Properties of Probabilistic Processes Combinators. In: Borgström, J., Crafa, S. (eds.) EXPRESS/SOS 2014. EPTCS, vol. 160, pp. 63–78 (2014)
15. Gebler, D., Tini, S.: SOS Specifications of Probabilistic Systems by Uniformly Continuous Operators. In: Aceto, L., Frutos-Escrig, D. (eds.) CONCUR 2015. LIPIcs, vol. 42, pp. 155–168. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik (2015)
16. Jonsson, B., Larsen, K.G., Yi, W.: Probabilistic Extensions of Process Algebras. In: Handbook of Process Algebra, pp. 685–710. Elsevier (2001)
17. Kermarrec, A.M., van Steen, M.: Gossiping in distributed systems. Operating Systems Review 41(5), 2–7 (2007)
18. Kwiatkowska, M., Norman, G., Parker, D.: Analysis of a gossip protocol in prism. SIGMETRICS Performance Evaluation Review 36(3), 17–22 (2008)
19. Lanotte, R., Merro, M.: Semantic analysis of gossip protocols for wireless sensor networks. In: Katoen, J.P., König, B. (eds.) CONCUR 2011. LNCS, vol. 6901, pp. 156–170. Springer (2011)
20. Lanotte, R., Merro, M.: A semantic theory of the internet of things. In: Lluch-Lafuente, A., Proença, J. (eds.) COORDINATION 2016. LNCS, vol. 9686, pp. 157–174. Springer (2016)
21. Lanotte, R., Merro, M.: A calculus of cyber-physical systems. In: Drewes, F., Martín-Vide, C. (eds.) LATA 2017. LNCS, vol. 10168, pp. 115–127 (2017)
22. Larsen, K.G., Skou, A.: Bisimulation through Probabilistic Testing. Information and Computation 94(1), 1–28 (1991)
23. Macedonio, D., Merro, M.: A semantic analysis of key management protocols for wireless sensor networks. Sci. Comput. Program. 81, 53–78 (2014)
24. Merro, M., Ballardin, F., Sibilio, E.: A timed calculus for wireless systems. Theoretical Computer Science 412(47), 6585–6611 (2011)
25. Segala, R., Lynch, N.: Probabilistic simulations for probabilistic processes. Nordic Journal of Computing 2, 250–273 (1995)
26. Villani, C.: Optimal transport, old and new. Springer (2008)