A semi-quantitative equivalence for abstracting from fast reactions

Vashti Galpin[1]   Jane Hillston[1,2]   Federica Ciocchetta

[1] LFCS, School of Informatics, University of Edinburgh

[2] CSBE, University of Edinburgh

Vashti.Galpin@ed.ac.uk, Jane.Hillston@ed.ac.uk

Semantic equivalences are used in process algebra to capture the notion of similar behaviour, and this paper proposes a semi-quantitative equivalence for a stochastic process algebra developed for biological modelling. We consider abstracting away from fast reactions as suggested by the Quasi-Steady-State Assumption. We define a fast-slow bisimilarity based on this idea. We also show congruence under an appropriate condition for the cooperation operator of Bio-PEPA. The condition requires that there is no synchronisation over fast actions, and this distinguishes fast-slow bisimilarity from weak bisimilarity. We also show congruence for an operator which extends the reactions available for a species. We characterise models for which it is only necessary to consider the matching of slow transitions and we illustrate the equivalence on two models of competitive inhibition.

1   Introduction

One of the features of process algebra is *behavioural or semantic equivalence* [29, 24] which determines if two processes act the same way. Furthermore, *congruence* is of interest where the behaviours of two equivalent systems are indistinguishable within any context, in the sense of combining systems using one or more operators of the process algebra. Notions of equivalence can be also useful in *systems biology*. For example, equivalences can be used to compare the behaviour of two systems or parts of them, and to show the consistency between different abstractions of the same system. Furthermore, if congruence holds then it may be possible to replace a (part of the) system with a smaller equivalent component without changing the behaviour, and thus reduce the state space.

We focus on Bio-PEPA [14], a process algebra defined for modelling and analysis of biochemical networks. Bio-PEPA models have a number of interpretations, including ordinary differential equations (ODEs) and stochastic simulation [21]. We develop our equivalence in the context of mapping a Bio-PEPA model to a *finite transition system* (which can be interpreted as a finite *continuous-time Markov chain* (CTMC)). Such models are called *Bio-PEPA models with levels* since we assume a maximum quantity for each species, and we stratify molecule counts or discretise concentrations into levels, resulting in a finite and tractable transition system, thus ameliorating the state space explosion problem.

Here we apply a traditional technique of process algebras, namely semantic equivalence to Bio-PEPA with levels. *Isomorphism* and *strong equivalence* (adapted from PEPA [24]) have both been defined for Bio-PEPA [14, 19], but both relations are very strong notions of equivalence and not able to capture biological behaviour of interest.

By contrast, our approach is *semi-quantitative* in that we consider *relative* rather than actual speeds of reactions, and use this to determine whether two models have similar behaviour by abstraction from the faster reactions. The more abstract model has fewer species and hence fewer parameters. This reduced model can be parameterised more straightforwardly when there is limited experimental data.

We define *fast-slow bisimilarity* for Bio-PEPA inspired by biology. The motivation comes from the *Quasi-Steady-State Assumption* (QSSA) [36] which may reduce systems of ODEs where there are large differences in reaction rates. This is achieved by abstracting from fast reactions, by assuming almost no change in the amount of intermediate products (effectively assuming a steady state for these products), obtaining fewer reactions and a smaller system of ODEs. The rates of reactions in the reduced system are no longer based on mass action but are defined in more complex manner which is determined in the derivation of the reduced system of ODEs. It is then possible to work with the reduced system as a model thereby requiring fitting of fewer parameters.

As with any semantic equivalence we investigate congruence. Fast-slow bisimilarity is shown to be a congruence with respect to cooperation under the condition that there are no fast actions that can occur between pairs of components. This restriction for congruence is not a significant limitation as it has been shown that introducing other fast reactions (beyond those that are considered by QSSA) involving the species to which QSSA is applied can drastically reduce the accuracy of the QSSA [35]. Fast-slow bisimilarity is similar in definition to an existing equivalence called *weak bisimilarity* where the behaviour of silent or invisible τ actions is abstracted away. However, the additional condition needed to show congruence distinguishes them. Using an existing technique [22], we are able to show that for certain reduced models, it is possible to work with a definition of bisimilarity which only considers slow reactions.

The rest of the paper is structured as follows. An introduction to QSSA follows, then Bio-PEPA is introduced and the definition of fast-slow equivalence is proposed. Congruence is proved for cooperation and the extension operator. Slow bisimulation is defined and conditions identified for which this is sufficient, followed by an example to illustrate our equivalence. Finally related and further work is discussed and concluding remarks are given.

2 Quasi-Steady-State Assumption

In this section, we consider an existing approach to model reduction which reduces the number of species to be considered and determines new reaction rates. First, consider a set of non-oscillating reactions. We can identify a set of reactants that are present before the reactions start, say Ξ and a set of products Ψ that are present once all reactions have completed. However, complexes may be created during the reactions – these are called *intermediate species*, and we use Υ for the set of these species. We will also call the species in $\Phi = \Xi \cup \Psi$ *non-intermediate*. Note that $\Upsilon \cap \Phi = \emptyset$ but we cannot assume that $\Xi \cap \Psi$ is empty since modifiers such as enzymes may appear in both.

In cellular systems, biochemical reactions can happen on very different time scales. There can be very frequent reactions (*fast reactions*) and less frequent reactions (*slow reactions*). In this case, we can apply the Quasi-Steady-State assumption [36] which is a time scale separation approach. We discuss other time scale separation/decomposition techniques in Section 8.

If fast reactions lead to the production of intermediate species, then the instantaneous rates of change of the intermediate species in the reaction are approximately equal to zero, with respect to the slow reactions, so they can be viewed as being at steady state. The pre-steady-state transient period before this happens is much shorter compared to the time taken for slower reactions, and since this period is typically of less interest, inaccuracies are less important.

Specifically, we assume species X_i, $i = 1, \ldots, n$ with $\Upsilon = \{X_{j_1}, \ldots, X_{j_m}\}$ the m intermediate species. The equation[1] $dX_{j_k}/dt = f_{j_k}(X)$ (where $f_{j_k}(X)$ stands for a mathematical expression describing the

---

[1] Here the concentration of the species X is represented by the variable X.

dynamics of X_{j_k} in terms of all species) is assumed to be approximately equal to zero, and therefore $f_{j_k}(X) \approx 0$. From these equations, it may be possible to derive an expression for intermediate species in terms of other species. These expressions can replace the variables representing intermediate species in the rate equations of other species. The resulting ODE system has fewer equations and terms (reactions). Therefore, the effect of the application of QSSA is to simplify the model complexity. This kind of approximation and resulting reduction is useful when there are many species and hence many ODEs; when systems of ODEs are stiff (they have widely different rates) and hence are difficult to solve numerically [9]; and when it is difficult experimentally to obtain rate parameters for intermediate species.

As an example, consider the reactions $S + E \xleftrightarrow{R_1, R_{-1}} SE \xrightarrow{R_2} P + E$ where S is the substrate, P is the product, E the enzyme, SE the intermediate substrate-enzyme compound. The double-headed arrow represents a reversible reaction, with the forward reaction named R_1 and the reverse reaction named R_{-1}. All reactions are described by mass-action kinetics with rate constants k_1, k_{-1}, k_2. It represents the *Michaelis-Menten mechanism for enzymatic catalysis* [36, 28]. The corresponding ODE system is

$$
\begin{aligned}
dS/dt &= -k_1 E \cdot S + k_{-1} SE & dE/dt &= -k_1 E \cdot S + k_{-1} SE + k_2 SE \\
dP/dt &= +k_2 SE & \text{and} \quad dSE/dt &= +k_1 E \cdot S - k_{-1} SE - k_2 SE
\end{aligned}
$$

When the first two reactions are assumed fast and the third slow, the intermediate species SE is considered to be at steady-state. We can say that $dSE/dt \approx 0$ and from this we obtain $SE = E_T \cdot S/(S + K_M)$, where $E_T = E + SE$ is the total enzyme in the system and constant, and $K_M = (k_{-1} + k_2)/k_1$ (*Michaelis-Menten constant*). Replacing SE with the expression above, we derive the ODE system $-dS/dt = dP/dt = (k_2 E_T \cdot S)/(S + K_M)$. We now have a single reaction $S + E \to P + E$ that abstracts the original reactions. This simplification is called the *Michaelis-Menten (MM) approximation* and it is valid under some assumptions, such $S_0 + K_m \gg E_T$ where S_0 is the initial quantity of S [37].

This approach provides an analogy for the development of our semantic equivalence which is presented in the next section. It leads us to partition reactions into fast and slow, and allows us to abstract away from the fast reactions. We now introduce the process algebra to which we will apply these concepts.

3   Bio-PEPA with levels

The syntax of Bio-PEPA with levels [14] is given by the grammar below. S defines sequential components which describe the behaviour of biochemical species, and P defines model components which combine the species components and from which we can understand the interactions between species.

$$
S ::= (\alpha, \kappa) \text{ op } S \mid S + S \mid C \qquad \text{op} ::= \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot \qquad P ::= P \underset{L}{\bowtie} P \mid S(l)
$$

In the term $(\alpha, \kappa) \text{ op } S$, α is an action or reaction name from $\mathscr{A}$, $\kappa \in \{1, 2, \ldots\}$ is the stoichiometric coefficient of the species and the prefix combinator op describes the role of the species in the reaction. The symbol $\downarrow$ is used for a reactant, $\uparrow$ a product, $\oplus$ an activator, $\ominus$ an inhibitor, and $\odot$ for a generic modifier. The operator $+$ provides the choice between two sequential components and species constants are defined by $C \overset{def}{=} S$. The process $P \underset{L}{\bowtie} Q$ denotes the synchronisation between two components P and Q and the set L specifies those reactions on which the components must synchronise. We use $P \underset{*}{\bowtie} Q$ to denote the case when all actions shared by P and Q are synchronised on. In the model component $S(l)$, the parameter $l \in \mathbb{N}$ represents the level of molecular count or concentration. The set of all Bio-PEPA species components is $\mathscr{S}$ and the set of all Bio-PEPA model components is $\mathscr{C}$.

We consider a constrained set of Bio-PEPA models to ensure well-behaved systems. We require that a species is a choice between reactions without any repeated actions and that there is only one species component for a species at model level, as described by the next definition.

Definition 3.1. *A Bio-PEPA sequential component C is* well-defined *if it has the form*

$$C \stackrel{\text{def}}{=} (\alpha_1, \kappa_1) \operatorname{op}_1 C + \ldots + (\alpha_n, \kappa_n) \operatorname{op}_n C \quad \text{written as} \quad C \stackrel{\text{def}}{=} \sum_{i=1}^{n} (\alpha_i, \kappa_i) \operatorname{op}_i C \quad \text{where } \alpha_i \neq \alpha_j \text{ for } i \neq j.$$

A model component P is well-defined *if it has the form* $P \stackrel{\text{def}}{=} C_1(l_1) \underset{L_1}{\bowtie} \ldots \underset{L_{p-1}}{\bowtie} C_p(l_p)$ *where each C_i is a well-defined sequential component, the elements of each L_j appear in P and if $i \neq j$ then $C_i \neq C_j$.*

Additionally, each model has an associated context collecting together information such as rates, compartments and parameters, as now defined.

Definition 3.2. *A well-defined Bio-PEPA system $\mathscr{P}$ is a six-tuple $\langle \mathscr{V}, \mathscr{N}, \mathscr{K}, \mathscr{F}, Comp, P \rangle$, where $\mathscr{V}$ is the set of compartments, $\mathscr{N}$ is the set of quantities describing each species, $\mathscr{K}$ is the set of parameters, $\mathscr{F}$ is the set of functional rates, Comp is the set of well-defined sequential components and P is a well-defined model component. $\mathscr{V}, \mathscr{N}, \mathscr{K}, \mathscr{F}, Comp$ are called the* context *of P.*

For details of the elements of the context and the definition of well-defined Bio-PEPA system, see [14, 13]. In this paper, we only consider single compartment models [2]. The levels for a species are obtained from information contained in $\mathscr{N}$ for that species. More specifically, for each species C, we assume a maximum[3] molecular count M_C, and a fixed step size H across all species to ensure conservation of mass during reactions involving multiple species. The maximum level for a species is determined by $N_C = \lceil M_C/H \rceil$. Thus, C has levels, $0, \ldots, N_C$, giving $N_C + 1$ levels in total.

The operational semantics for Bio-PEPA systems with levels is given in Figure 1 where N_S is the maximum level for the species S. These operational semantics define two labelled transition systems. The enzyme, inhibitor and general modifier prefixes are used in reactions that are not modelled as bimolecular reactions with mass action kinetics, hence the semantics for these prefixes reflect the fact that the species is not consumed in the reaction and the level remains the same.

The rules with lowercase letters derive a relation where transitions are labelled with a reaction name and a string collecting information about each species involved in the reaction consisting of the species name, its role in the reaction, its stoichiometric coefficient for the reaction, and its current level. This information is then used in the rule `Final` which includes the context of the model to determine the rate of the reaction and generates the *stochastic relation*[4] from which a CTMC can be obtained. In this paper, we focus on the capability relation as we use this in our definition of equivalence.

Definition 3.3. *A capability label is defined as $\theta = (\alpha, w)$ with $\alpha \in \mathscr{A}$ and the list w defined by the grammar $w ::= [S:\operatorname{op}(l, \kappa)] \mid w::w$ where $S \in \mathscr{S}, l \in \mathbb{N}, l \geq 0$ the level, $\kappa \in \mathbb{N}, \kappa \geq 1$ the stoichiometric coefficient, and :: is list concatenation. The set of all such capability labels is Θ.*

Definition 3.4. *Given a Bio-PEPA model, the* capability relation $\to_c \subseteq \mathscr{C} \times \Theta \times \mathscr{C}$ *is the smallest relation defined by the first nine rules in Figure 1. An element of the transition system is written $P \xrightarrow{(\alpha, w)}_c P'$.*

The string w is defined as a list [14] but the order of elements is not important so it can be viewed as a multiset. For well-defined systems, w is a set [19] and we will treat it as such in the sequel.

---

[2]For a presentation of Bio-PEPA with locations see [12].

[3]This is reasonable since cells and other biological compartments have constrained volumes.

[4]For more details on the stochastic relation defined by `Final`, see [14, 13]. Briefly, $r_\alpha[w, \mathscr{N}, \mathscr{K}] = f_\alpha[w, \mathscr{N}, \mathscr{K}]/H \in (0, \infty)$ where f_α is the functional rate for the reaction α from $\mathscr{F}$ and H is the step size. $\mathscr{N}$ provides species information and $\mathscr{K}$ provides constants.

prefixReac
$$\frac{}{(\alpha,\kappa){\downarrow}S(l) \xrightarrow{(\alpha,[S:\downarrow(l,\kappa)])}_c S(l-\kappa)} \quad \kappa \leq l \leq N_S$$

prefixProd
$$\frac{}{(\alpha,\kappa){\uparrow}S(l) \xrightarrow{(\alpha,[S:\uparrow(l,\kappa)])}_c S(l+\kappa)} \quad 0 \leq l \leq N_S - \kappa$$

prefixMod
$$\frac{}{(\alpha,\kappa)\ \mathrm{op}\ S(l) \xrightarrow{(\alpha,[S:\mathrm{op}(l,\kappa)])}_c S(l)} \quad \begin{array}{l} \kappa \leq l \leq N_S \text{ if op} = \oplus \\ 0 \leq l \leq N_S \text{ if op} \in \{\ominus, \odot\} \end{array}$$

choice1
$$\frac{S_1(l) \xrightarrow{(\alpha,w)}_c S_1'(l')}{(S_1+S_2)(l) \xrightarrow{(\alpha,w)}_c S_1'(l')}$$

choice2
$$\frac{S_2(l) \xrightarrow{(\alpha,w)}_c S_2'(l')}{(S_1+S_2)(l) \xrightarrow{(\alpha,w)}_c S_2'(l')}$$

coop1
$$\frac{P_1 \xrightarrow{(\alpha,w)}_c P_1'}{P_1 \underset{L}{\bowtie} P_2 \xrightarrow{(\alpha,w)}_c P_1' \underset{L}{\bowtie} P_2} \quad \alpha \notin L$$

coop2
$$\frac{P_2 \xrightarrow{(\alpha,w)}_c P_2'}{P_1 \underset{L}{\bowtie} P_2 \xrightarrow{(\alpha,w)}_c P_1 \underset{L}{\bowtie} P_2'} \quad \alpha \notin L$$

coop3
$$\frac{P_1 \xrightarrow{(\alpha,w_1)}_c P_1' \quad P_2 \xrightarrow{(\alpha,w_2)}_c P_2'}{P_1 \underset{L}{\bowtie} P_2 \xrightarrow{(\alpha,w_1::w_2)}_c P_1' \underset{L}{\bowtie} P_2'} \quad \alpha \in L$$

constant
$$\frac{S(l) \xrightarrow{(\alpha,[S:\ \mathrm{op}\ (l,\kappa)])}_c S'(l')}{C(l) \xrightarrow{(\alpha,[C:\ \mathrm{op}\ (l,\kappa)])}_c S'(l')} \quad C \stackrel{def}{=} S$$

Final
$$\frac{P \xrightarrow{(\alpha,w)}_c P'}{\langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}, Comp, P \rangle \xrightarrow{(\alpha,r_\alpha[w,\mathcal{N},\mathcal{K}])}_s \langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}, Comp, P' \rangle}$$

Figure 1: Operational semantics of Bio-PEPA

Definition 3.5. *The* derivative set $ds(P)$ *is the smallest set such that* $P \in ds(P)$ *and if* $P' \in ds(P)$ *and* $P' \xrightarrow{(\alpha,w)}_c P''$ *then* $P'' \in ds(P)$. $P' \in ds(P)$ *is a* derivative *of P.*

This section has defined Bio-PEPA syntax and semantics for Bio-PEPA systems with levels. We assume well-defined Bio-PEPA systems with levels for the remainder of the paper. Moreover, we assume that in well-defined Bio-PEPA systems all shared actions are synchronised over and hence we use the aforementioned notation $\underset{*}{\bowtie}$ for cooperation. The next section considers the equivalence we develop.

4 Fast-slow bisimilarity

Our basis for developing the equivalence is the QSSA where intermediate species at steady state can be approximated. As defined in Section 2, we have a set of intermediate species Υ and non-intermediate species Φ with $\Upsilon \cap \Phi = \emptyset$. When comparing models, we need to ensure that we exclude intermediate species in the comparison. Therefore we define a function that transforms the set w by removing all intermediate species in Υ, and leaving certain species in $\Delta \subseteq \Phi$. Typically, these species will be those that appear in transitions for slow reactions in both models with the same role in both reactions. Let $w_\Delta = \{C:\mathrm{op}\,(l,\kappa) \in w \mid C \in \Delta\}$ and note that $(w_1::w_2)_\Delta = (w_1)_\Delta::(w_2)_\Delta$.

Since QSSA is based on relative reaction rates, we assume that each reaction can be described as fast or slow, leading to a partition of the set $\mathscr{A}$ into the set of slow reactions $\mathscr{A}_s$ and the set of fast reactions $\mathscr{A}_f$. For convenience, we introduce new transitions.

Definition 4.1. *For well-defined Bio-PEPA models* P, P'.

- *If* $P \xrightarrow{(\alpha,w)}_c P'$ *and* $\alpha \in \mathscr{A}_f$ *then* $P \twoheadrightarrow P'$.   • *If* $P \xrightarrow{(\alpha,w)}_c P'$ *and* $\alpha \in \mathscr{A}_s$ *then* $P \xrightarrow{\alpha,w_\Delta} P'$.
- *If* $P(\twoheadrightarrow)^* P'$ *then* $P \Rightarrow P'$.   • *If* $P(\twoheadrightarrow)^* \xrightarrow{\alpha,w_\Delta} (\twoheadrightarrow)^* P'$ *then* $P \xRightarrow{\alpha,w_\Delta} P'$.

These new transitions consider the reaction names, the reaction speeds and certain non-intermediate species involved in the reaction, hence the equivalence defined will be semi-quantitative in nature. We now define a new bisimulation based on fast and slow actions.

Definition 4.2. *A symmetric relation* $\mathscr{R}$ *over* $\mathscr{C} \times \mathscr{C}$ *is a* fast-slow bisimulation *for* $\mathscr{A}_f$ *if* $(P,Q) \in \mathscr{R}$ *implies that*

- *for all* $\alpha \in \mathscr{A}_s$ *whenever* $P \xrightarrow{\alpha,w_\Delta} P'$ *there exists* Q' *with* $Q \xRightarrow{\alpha,w_\Delta} Q'$ *and* $(P',Q') \in \mathscr{R}$, *and*
- *whenever* $P \twoheadrightarrow P'$ *there exists* Q' *with* $Q \Rightarrow Q'$ *and* $(P',Q') \in \mathscr{R}$

Here $\twoheadrightarrow$ plays a similar role to $\xrightarrow{\tau}$ in Milner's definition of weak bisimulation [29], and hence these are similar notions. Some results for weak bisimulation may hold for fast-slow bisimulation but there are limits on this, particularly for proofs based on transition derivations. For example, when showing congruence where one works with transition derivations, the difference between $\twoheadrightarrow$ and $\xrightarrow{\tau}$ is apparent – this will be discussed in more detail later when congruence with respect to the cooperation is proved. We can now define a notion of fast-slow bisimilarity with respect to a given set of fast actions.

Definition 4.3. P *and* Q *are* fast-slow bisimilar *for* $\mathscr{A}_f$ $(P \approx_{\mathscr{A}_f} Q)$ *if there exists a fast-slow bisimulation for* $\mathscr{A}_f$, $\mathscr{R}$ *such that* $(P,Q) \in \mathscr{R}$.

Then $\approx_{\mathscr{A}_f}$ is the largest fast-slow bisimulation for $\mathscr{A}_f$. Now that we have a definition, we wish to show that it is useful by proving congruence for operators of interest, and by considering an example.

5   Congruence

When a semantic equivalence captures the notion of same behaviour we can reason about pairs of systems acting the same. However, if we show that a semantic equivalence is a congruence with respect to an operator of the process algebra, then we know that we can build new systems with equivalent behaviour using that operator. We start by considering the cooperation operator as it would be useful to know that we can combine fast-slow bisimilar systems.

To ensure congruence, it is not possible for fast actions to appear on both sides of the cooperation operator. This makes sense since this equivalence abstracts away from the details of the fast reactions, and it is not possible to know if the two models have abstracted the same reactions. Moreover, recent assessment of Michaelis-Menten approximation has shown that the QSSA does not hold if there are other fast reactions (apart from those to which the QSSA is applied) involving the species that are part of the Michaelis-Menten reactions [35].

Theorem 5.1. *If* $P_1 \approx_{\mathscr{A}_f} P_2$ *then* $P_1 \bowtie_* Q \approx_{\mathscr{A}_f} P_2 \bowtie_* Q$ *and* $Q \bowtie_* P_1 \approx_{\mathscr{A}_f} Q \bowtie_* P_2$ *provided that no action in* $\mathscr{A}_f$ *appears in both* Q *and* P_1 *or in both* Q *and* P_2.

Proof. Let $\mathscr{R} = \{(P_1' \bowtie_* Q', P_2' \bowtie_* Q') \mid P_1' \approx_{\mathscr{A}_f} P_2'\}$. Consider a transition $P_1' \bowtie_* Q' \xrightarrow{\alpha,w_\Delta} R$ which is obtained from $P_1' \bowtie_* Q' \xrightarrow{(\alpha,w)}_c R$ since $\alpha \in \mathscr{A}_s$. There are three cases and we prove the most complex here. Assume $P_1' \xrightarrow{(\alpha,w_1)}_c P_1''$, $Q' \xrightarrow{(\alpha,w_2)}_c Q''$ and $w = w_1 :: w_2$ then R is $P_1'' \bowtie_* Q''$. Since $P_1' \approx_{\mathscr{A}_f} P_2'$, $P_2' \xRightarrow{\alpha,(w_1)_\Delta} P_2''$ and hence $P_2' \xrightarrow{(\beta_1,v_1)}_c \cdots \xrightarrow{(\beta_n,v_n)}_c P_3' \xrightarrow{(\alpha,w_1)}_c P_3'' \xrightarrow{(\gamma_1,u_1)}_c \cdots \xrightarrow{(\gamma_m,u_m)}_c P_2''$ for the actions $\beta_1,\ldots,\beta_n,\gamma_1,\ldots,\gamma_m \in \mathscr{A}_f$. From this, we can derive $P_2' \bowtie_* Q' \xrightarrow{(\beta_1,v_1)}_c \cdots \xrightarrow{(\beta_n,v_n)}_c P_3' \bowtie_* Q' \xrightarrow{(\alpha,w_1::w_2)}_c$

$P_3'' \bowtie_* Q'' \xrightarrow{(\gamma_1,u_1)}_c \cdots \xrightarrow{(\gamma_m,u_m)}_c P_2'' \bowtie_* Q''$, hence we have $P_2' \bowtie_* Q' \xrightarrow{\alpha,(w_1::w_2)_\Delta} P_2'' \bowtie_* Q''$ as required with $(P_1'' \bowtie_* Q'', P_2'' \bowtie_* Q'') \in \mathscr{R}$.

Next, consider a transition $P_1' \bowtie_* Q' \rightarrow R$. Hence there exists $\beta \in \mathscr{A}_f$ such that $P_1' \bowtie_* Q' \xrightarrow{(\beta,v)}_c R$. There are two cases since cooperation over an action in $\mathscr{A}_f$ is excluded by the condition. We show the case where $Q' \xrightarrow{(\beta,v)}_c Q''$ and R is $P_1' \bowtie_* Q''$. Then the transition $P_2' \bowtie_* Q' \xrightarrow{(\beta,v)}_c P_2' \bowtie_* Q''$ can be derived and hence $P_2' \bowtie_* Q' \Rightarrow P_2' \bowtie_* Q''$ with $(P_1' \bowtie_* Q'', P_2' \bowtie_* Q'') \in \mathscr{R}$ as required. $\qquad\square$

To see why the sharing of fast actions must be prohibited, consider $S_1 \stackrel{def}{=} (\alpha,2) \uparrow S_1 + (\gamma,2) \downarrow S_1$ and $S_2 \stackrel{def}{=} (\beta,2) \uparrow S_2 + (\gamma,2) \downarrow S_2$. Clearly $S_1(0) \approx_{\{\alpha,\beta\}} S_2(0)$. Considering a third species $S \stackrel{def}{=} (\alpha,1) \downarrow S$, it is not the case that $S_1(0) \bowtie_* S(1) \approx_{\{\alpha,\beta\}} S_2(0) \bowtie_* S(1)$ since these systems have very different behaviours because there are no γ reactions in the first systems and there are repeating γ reactions in the second. To prevent this difference, S could be modified to require that it perform all fast reactions giving $S \stackrel{def}{=} (\alpha,1) \downarrow S + (\beta,1) \downarrow S$ but it is not clear how to generalise this beyond species. This counter-example for the condition in the theorem demonstrates that the definition of fast-slow bisimulation does differ from that of weak bisimulation. Here we abstract away from fast reaction names on a transition, whereas with weak bisimulation, transitions with the named action τ are treated abstractly.

5.1 The species extension operator

Since the focus is well-defined Bio-PEPA models, it is not useful to consider the operators for sequential agents individually. However, we do sometimes want to extend species' ability to be involved in reactions, and we now define an operator that permits this to happen and show that we have congruence for this operator under certain conditions.

When we build the models of biological systems, we may have no knowledge of what other reactions the species could be involved in. For example, considering the product of some sequence of reactions, we can imagine various scenarios in which we would want the product to be able to react with new species added by cooperation. It is difficult to do this at model level but we can consider it at species level by defining an appropriate extension operator [19].

Definition 5.1. *Given two well-defined species A and B such that the reaction names of A are disjoint from those of B, define $A\{B\}$, the* extension *of A by B as*

$$A\{B\} = \sum_{i=1}^{n} (\alpha_i, \kappa_i)\mathrm{op}_i A\{B\} + \sum_{j=1}^{m} (\beta_j, \lambda_j)\mathrm{op}_j A\{B\} \quad where \quad A = \sum_{i=1}^{n} (\alpha_i, \kappa_i)\mathrm{op}_i A, \quad B = \sum_{j=1}^{m} (\beta_j, \lambda_j)\mathrm{op}_j B.$$

This permits A to take on additional reaction capabilities, specifically those of B. $A\{B\}$ is well-defined since there are no repeated reaction names and because A and B are well-defined. Note that $A\{B\}$ and $B\{A\}$ are isomorphic since their transition systems are structurally identical with matching actions. The next theorem shows how species can be augmented with ways to participate in new reactions in a way which preserves fast-slow bisimulation.

Theorem 5.2. *Let $C_1(l) \approx_{\mathscr{A}_f} C_2(l)$ for sequential Bio-PEPA components C_1 and C_2 for all $0 \leq l \leq N_{C_1} = N_{C_2}$ and let C be a well-defined species with reaction names disjoint from those in C_1 and C_2. Then $C_1\{C\}(l) \approx_{\mathscr{A}_f} C_2\{C\}(l)$ and $C\{C_1\}(l) \approx_{\mathscr{A}_f} C\{C_2\}(l)$.*

Proof. Consider a transition $C_1\{C\}(l) \xrightarrow{\alpha,w_\Delta} C_1\{C\}(l')$ for $\alpha \in \mathscr{A}_s$ then $C_1\{C\}(l) \xrightarrow{(\alpha,w)}_c C_1\{C\}(l')$. If α appears in C, then $C_2\{C\}(l) \xrightarrow{(\alpha,w)}_c C_2\{C\}(l')$ and $C_2\{C\}(l) \xrightarrow{\alpha,w_\Delta} C_2\{C\}(l')$.

If α appears in C_1 and C_2 then $C_1(l) \xrightarrow{\alpha,w_\Delta} C_1(l')$ and since $C_1(l) \approx_{\mathscr{A}_f} C_2(l)$, then $C_2(l) \xrightarrow{\alpha,w_\Delta} C_2(l'')$. This then gives $C_2(l) \xrightarrow{(\beta_1,v_1)}_c \cdots \xrightarrow{(\beta_n,v_n)}_c C_2(m) \xrightarrow{(\alpha,w)}_c C_2(m') \xrightarrow{(\gamma_1,u_1)}_c \cdots \xrightarrow{(\gamma_m,u_m)}_c C_2(l'')$ for the actions $\beta_1, \ldots, \beta_n, \gamma_1, \ldots, \gamma_m \in \mathscr{A}_f$. $C_2\{C\}(l)$ can perform the same actions hence $C_2\{C\}(l) \xrightarrow{\alpha} C_2\{C\}(l'')$.

Next, consider $C_1\{C\}(l) \twoheadrightarrow C_1\{C\}(l')$. There exists $\beta \in \mathscr{A}_f$ such that $C_1\{C\}(l) \xrightarrow{(\beta,v)}_c C_1\{C\}(l')$. If β appears in C then as above, there is a matching transition in $C_2\{C\}(l)$. If β appears in C_1 since $C_1(l) \approx_{\mathscr{A}_f} C_2(l)$, then $C_2(l) \Rightarrow C_2(l'')$ which gives $C_2(l) \xrightarrow{(\beta_1,v_1)}_c \cdots \xrightarrow{(\beta_n,v_n)}_c C_2(l'')$ for $\beta_1,\ldots,\beta_n \in \mathscr{A}_f$. $C_2\{C\}(l)$ can perform these actions, hence $C_2\{C\}(l) \Rightarrow C_2\{C\}(l'')$ as required. $\square$

Let $C_1 \stackrel{def}{=} (\alpha,1)\uparrow C_1$ and $C_2 \stackrel{def}{=} (\alpha,1)\uparrow C_2 + (\beta,1)\oplus C_2$ with shared maximum level, then $C_1(l) \approx_{\{\beta\}} C_2(l)$. For any sequential component C with different reaction names to α and β and the same maximum level, the congruence result applies.

6   Slow bisimilarity

Checking for fast-slow bisimilarity requires that all reactions must be considered. We now define an equivalence over just the slow reactions. If we can identify conditions under which this equivalence implies fast-slow bisimilarity, then whenever we have models that satisfy those conditions, we need only check the slow reactions to prove that a relation is a fast-slow bisimulation. This section introduces such an equivalence and conditions. In the next section when we consider competitive inhibition, we will illustrate how compact our proofs are. First we define the new equivalence. As before, it is assumed that $\mathscr{A}_s$ and $\mathscr{A}_f$ partition $\mathscr{A}$.

Definition 6.1. *A symmetric relation $\mathscr{R}$ over $\mathscr{C} \times \mathscr{C}$ is a* slow bisimulation for $\mathscr{A}_s$ *if* $(P,Q) \in \mathscr{R}$ *implies that for all* $\alpha \in \mathscr{A}_s$ *whenever*

- *$P \xrightarrow{\alpha,w_\Delta} P'$ there exists Q' with $Q \xRightarrow{\alpha,w_\Delta} Q'$ and $(P',Q') \in \mathscr{R}$*

P and Q are slow bisimilar for $\mathscr{A}_s$ *if there exists a slow bisimulation for $\mathscr{A}_s$, $\mathscr{R}$ such that $(P,Q) \in \mathscr{R}$.*

We now consider when this can be applied using a technique that allow variables to be classified by what reactions affect them.

6.1   Variable classification

We present an existing technique that allows the identification of slow variables (those that are only affected by slow reactions) and fast variables (those that are affected by fast and slow reactions) [22]. Note that variables are not the same as species since a variable can either be an individual species or a linear combinations of species.

A set of reactions can be expressed as a stoichiometry matrix $\mathbf{S}$ which has m columns, one for each reaction and n rows, one for each species. $S_{i,j}$ describes the stoichiometry of species X_i with respect to reaction R_j.

A stoichiometry matrix can be transformed into a matrix of the same size with the form described in Figure 2 [22]. As mentioned above, the variables that are associated with the rows of $\mathbf{Q}$ are linear combinations of the original species variables and hence when given values for the new variables, it is possible to establish values for the species.

The top row of submatrices in the figure are zeroes since these represent conserved variables and re-actions do not affect them. $\mathbf{Q}_{ss}$ has size $n_s \times m_s$ and captures the stoichiometry of slow reactions for slow variables. The other submatrix in its row is zero since slow variables are not affected by fast reactions. The last row of $\mathbf{Q}$ consists of an $n_f \times m_s$ matrix and an $n_f \times m_f$ matrix describing the stoichiometry of slow and fast reactions with respect to fast variables.

$$\mathbf{Q} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{Q}_{ss} & \mathbf{0} \\ \mathbf{Q}_{fs} & \mathbf{Q}_{ff} \end{bmatrix}$$

- Columns represent reactions $R_1, \ldots R_{m_s}, R'_1 \ldots R'_{m_f}$ where the R_i are m_s slow reactions, the R'_j are m_f fast reactions, and $m_s + m_f = m$
- Rows represent variables $X_1, \ldots, X_{n_c}, X'_1, \ldots X'_{n_s}, X''_1, \ldots X''_{n_f}$ where the X_i are conserved variables, the X'_j are slow variables, the X''_k are fast variables and $n_c + n_s + n_f = n$

Figure 2: Matrix transformation [22].

For a given ordering of reactions and variables, $\mathbf{Q}$ is unique. However, for different orderings, an equally valid but different $\mathbf{Q}$ may be obtained. If there are no slow variables then this technique cannot be used.

For reasons of space, it is not possible to fully describe the matrix transformation defined in [22]. The idea is based around invariants. These are variables whose values are not changed by the dynamics of the model. First, invariants (conserved variables) of the model are identified. Then, by removing the slow reactions from the reaction equations, it is possible to find slow variables (invariants when slow reactions are removed), if any. Then sufficient fast variables must be identified so that there are n variables in total. Each new variable must be linearly independent of the other new variables, and the new variables are linear combinations of the original species variables. This process is illustrated in the example section.

6.2 Application to Bio-PEPA

Given a Bio-PEPA model, we can construct its stoichiometry matrix from the species component definitions. Using the technique described above we can identify invariants, slow and fast variables[5].

Note that a well-defined Bio-PEPA model only differs from its derivatives in terms of the levels of each species, hence models can be represented as vectors where each element represents the level of a species. See Figure 3 for an example.

A model's transition system, the capability relation, is then defined over states that are given in vector form $(v_1, \ldots, v_p)$ for p species. These states can be transformed to $(s_1, \ldots, s_{n_s}, f_1, \ldots, f_{n_f})$ where $n_s + n_f = p - n_c$, producing a new transition system where the transitions are unchanged and the states are defined with respect to the values of the new variables, specifically the slow and fast variables. Conserved variable values are not included in the new states since their values are fixed. The new states contain the same information as the original states, and they therefore stay unique. It is not possible for two states in the original transition system to collapse into one state in the new transition system. We can conclude that the transition systems are isomorphic, meaning that there is a bijection between states, and transitions are preserved with the same labels.

We now identify conditions that allow us to show when a slow bisimulation is a fast-slow bisimulation using variable classification. We restrict ourselves to the case of equivalence between a model which has conserved, slow and fast variables and a model that has conserved variables, slow variables which are the same as those in the first model, and no fast variables. We also require that slow variables are individual species. Extending the result to more general cases is further work.

Proposition 6.1. *Consider two Bio-PEPA models P_i for $i = 1,2$ where Δ_i contains exactly the slow species of the model, such that Δ_1 and Δ_2 have the same species. Let $\mathcal{R}$ be a relation over Bio-PEPA models such that for all $\big((s_1,\ldots,s_{n_s},f_1,\ldots,f_{n_f}),(s'_1,\ldots,s'_{n_s})\big) \in \mathcal{R}$, the s_i and s'_i are values for all slow variables and the f_i are the values for fast variables If $s_i = s'_i$ for $i \in \{1,\ldots n_s\}$ and $\mathcal{R}$ is a slow bisimulation for $\mathscr{A}_s$, then $\mathcal{R}$ is a fast-slow bisimulation for $\mathscr{A}_f$.*

Proof. Let $\mathcal{R}$ be a slow bisimulation with the required condition. Hence we need to consider fast actions only. Let $\big((s_1,\ldots,s_{n_s},f_1,\ldots,f_{n_f}),(s_1,\ldots,s_{n_s})\big) \in \mathcal{R}$ and consider the fast transition such that $(s_1,\ldots,s_{n_s},f_1,\ldots,f_{n_f}) \twoheadrightarrow (s_1,\ldots,s_{n_s},f'_1,\ldots,f'_{n_f})$. We know that $(s_1,\ldots,s_{n_s}) \Rightarrow (s_1,\ldots,s_{n_s})$ and also that $\big((s_1,\ldots,s_{n_s},f'_1,\ldots,f'_{n_f}),(s_1,\ldots,s_{n_s})\big) \in \mathcal{R}$. There are no fast actions from $(s_1,\ldots,s_{n_s})$ to consider. $\square$

Given two Bio-PEPA models, the general technique can be summarised as follows.

1. Classify the variables in each model. Check that one model only has slow variables and that the slow variables are species and the same between models. If not, try different variable orderings.

2. Transform the transition systems of both models as described above.

3. Define a relation over the transformed transition systems of the two models where slow variables have the same value in each pair in the relation.

4. Check that this relation is a slow bisimulation, and use Proposition 6.1 to show that that it is a fast-slow bisimulation.

5. Since the transformation has provided an isomorphic transition system, the original models are fast-slow bisimilar.

7   Competitive inhibition

We now consider an example where there are significantly different rates and hence a suitable test case for fast-slow bisimulation. It is an example of competitive inhibition [36] where an inhibitor is introduced, giving the reactions $S + EI \longleftrightarrow S + E + I \longleftrightarrow SE + I \rightarrow P + E + I$. Here, the first reversible reaction describes how the enzyme and inhibitor can bind together to form a compound. The second reversible reaction shows how the substrate and enzyme bind together to form a compound from which the product can be obtained. The binding of the inhibitor and enzyme competes with the binding of the enzyme and substrate since when the enzyme is bound to the inhibitor it is not available for the reaction with the substrate and hence reduces the amount of product that can be produced. Then $\Xi = \{S,E,I\}$, $\Psi = \{P,E,I\}$ and $\Upsilon = \{EI,SE\}$ since EI and SE are the intermediate species (as defined in Section 2) created by these reactions. Because of the explicit representations of the inhibitor and enzyme, and their associated intermediates, we choose to model the basic bimolecular reactions with mass actions kinetics.

These reactions can be expressed in Bio-PEPA as follows, where α_1 and α_{-1} are the reactions involving enzyme and inhibitor, β_1 and β_{-1} are the reactions involving substrate and enzyme, and γ is the reaction that produces the product.

$$S \stackrel{def}{=} (\beta_1,1){\downarrow}S + (\beta_{-1},1){\uparrow}S \qquad\qquad P \stackrel{def}{=} (\gamma,1){\uparrow}P \qquad I \stackrel{def}{=} (\alpha_1,1){\uparrow}I + (\alpha_{-1},1){\downarrow}I$$

$$EI \stackrel{def}{=} (\alpha_1,1){\downarrow}EI + (\alpha_{-1},1){\uparrow}EI \qquad SE \stackrel{def}{=} (\beta_1,1){\uparrow}SE + (\beta_{-1},1){\downarrow}SE + (\gamma,1){\downarrow}SE$$

$$E \stackrel{def}{=} (\alpha_1,1){\uparrow}E + (\alpha_{-1},1){\downarrow}E + (\beta_1,1){\downarrow}E + (\beta_{-1},1){\uparrow}E + (\gamma,1){\uparrow}E$$

$$Sys \stackrel{def}{=} S(l_S) \underset{*}{\bowtie} E(l_E) \underset{*}{\bowtie} I(l_I) \underset{*}{\bowtie} P(l_P) \underset{*}{\bowtie} EI(l_{EI}) \underset{*}{\bowtie} SE(l_{SE})$$

Here, based on biological understanding, we set $\{\alpha_1,\alpha_{-1},\beta_1,\beta_{-1}\} = \mathscr{A}_f$, namely that these are the fast reactions, and that $\gamma \in \mathscr{A}_s$.
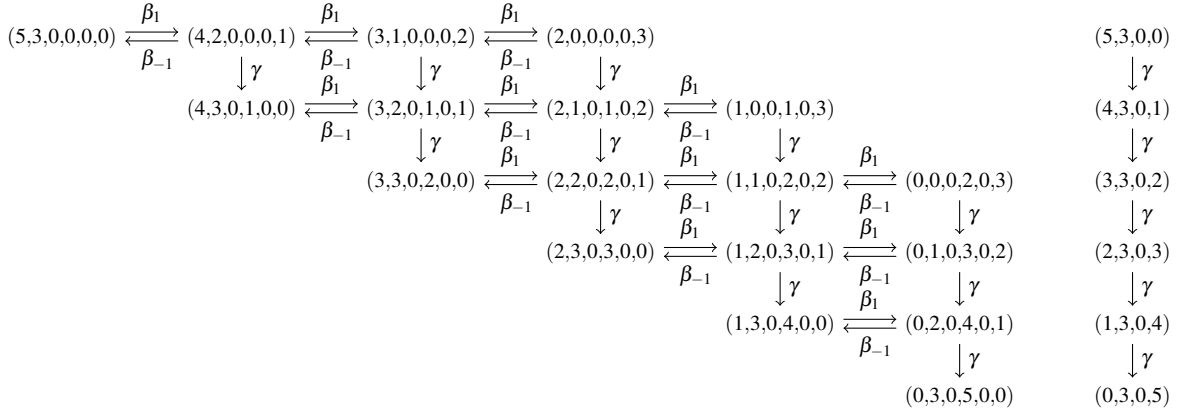
Figure 3: Transition system for *Sys* and *Sys'* for $n = 5$ and $m = 3$ with no inhibitor present (only reaction names appear on transitions).

We wish to show that this is fast-slow bisimilar to the simpler Bio-PEPA model defined as follows. In this model, only a single reaction is modelled and this reaction has a rate which takes into account the amount of enzyme and inhibitor present. This reaction is considered to be a slow reaction. Since this reaction is not based on mass actions kinetics, the inhibitor and enzyme prefix operators are used. The reaction is named γ as it produces P, as in the previous model. The use of primes on species and model names is a syntactic convenience to distinguish different species and model components. However, later on we will view P and P' as the same when we consider w_Δ.

$$S' \stackrel{def}{=} (\gamma, 1) \downarrow S' \qquad E' \stackrel{def}{=} (\gamma, 1) \oplus E' \qquad I' \stackrel{def}{=} (\gamma, 1) \ominus I' \qquad P' \stackrel{def}{=} (\gamma, 1) \uparrow P'$$

$$Sys' \stackrel{def}{=} S'(l_{S'}) \bowtie_* E'(l_{E'}) \bowtie_* I'(l_{I'}) \bowtie_* P'(l_{P'})$$

From the equations, it is clear that for a starting level of the substrate S (or S') given by $l_S = n$ it is not possible to reach more than n levels of P (alternatively P') if the starting level of P is set to $l_P = 0$. This agrees with the biological understanding that these reactions represent a transformation of S to P through a number of steps.

We assume neither of the intermediates nor the product are present at the start in the more complex model as is standard [37]. Hence, for a starting level of m of the enzyme, it is not possible to have more than m levels of the substrate-enzyme compound, and for a starting level of m of the enzyme and p of the inhibitor, it is not possible to have more than $\min\{m, p\}$ levels of the enzyme-inhibitor compound.

As mentioned above, a well-defined Bio-PEPA model only differs from its derivatives in the levels of the species and models and derivatives can be expressed in numeric vector form. For example, for *Sys* the vector $(2, 0, 3, 1, 0, 4)$ describes the model with 2 levels of S, none of E, 3 of I, 1 of the product P, none of the compound EI and 4 of the compound SE.

7.1 Constructing the bisimulation

Under the initial species levels described above, there are four cases of interest: only substrate present, substrate and enzyme present, substrate and inhibitor present, and substrate, enzyme and inhibitor present. These can be considered in one relation over the two models with starting vectors $(n, m, p, 0, 0, 0)$ (using the ordering $(l_S, l_E, l_I, l_P, l_{EI}, l_{SE})$) and $(n, m, p, 0)$ (using the ordering $(l_{S'}, l_{E'}, l_{I'}, l_{P'})$) with $n > 0$ and $m, p \geq 0$. Figure 3 illustrates the case when $n = 5$, $m = 3$ and $p = 0$. This case with no inhibitor represents an instance of the standard Michaelis-Menten mechanism [36] as discussed in Section 2.

	Sys: new variables				
X_{S_T}	$=$ $S+SE+P$	$=$ S_0	$=$ n	conserved
X_{E_T}	$=$ $E+EI+SE$	$=$ E_0	$=$ m	conserved
X_{I_T}	$=$ $EI+I$	$=$ I_0	$=$ p	conserved
X_P	$=$ P		$=$ k	slow
X_{EI}	$=$ EI		$=$ l	fast
X_{SE}	$=$ SE		$=$ j	fast

new state: (P, EI, SE)

	Sys$'$: new variables			
$X_{S'_T}$	$=$ $S'+P'$	$=$ S_0	$=$ n	conserved
$X_{E'}$	$=$ E'_0		$=$ m	conserved
$X_{I'}$	$=$ I'_0		$=$ p	conserved
$X_{P'}$	$=$ P'		$=$ k	slow

new state: (P)

Figure 4: Identification of conserved, fast and slow variables

To define fast-slow bisimulation, we must determine which non-intermediate species are in the set Δ. The label on the transition of a γ reaction in *Sys* is (γ, w) where $w = \{P{:}\uparrow(1, i_1), SE{:}\downarrow(1, i_2), E{:}\uparrow(1, i_3)\}$. For a γ in *Sys$'$*, the set is $\{P'{:}\uparrow(1, j_1), S'{:}\downarrow(1, j_2), E'{:}\oplus(1, j_3), I'{:}\ominus(1, j_4)\}$. We only want to compare species that appear in both sets and that have the same role, hence we let $\Delta = \{P\} = \{P'\}$. We will show below that the product is also the slow variable of both systems, illustrating another way to determine Δ.

Next define the relation $\mathscr{R}$ as
$$\big\{ ((n-(k+j), m-(j+l), p-l, k, l, j), (n-k, m, p, k)) \mid 0 \le k \le n, 0 \le j \le \min\{m, n-k\}, 0 \le l \le p, j+l \le m \big\}.$$
This captures the idea suggested by Figure 3 that states with the same level of product are those that should be paired in $\mathscr{R}$. We now show that $\mathscr{R}$ is a fast-slow bisimulation for $\mathscr{A}_f$ using the approach given in the previous section. Figure 4 provides the new variables for each model. Here, variables with subscript 0 indicate initial values for those species. First three invariants are identified, then we consider just the fast transitions and this allows us to determine which species are not affected by the fast transitions. P is not affected and neither is $S+SE$. Since these are not linearly independent (due to the first invariant), we need to choose one of them, and we choose P since it is a single species. There are no other linearly independent slow variables so we need to find two fast variables that are linearly independent from each other and the four defined variables. EI and SE are suitable candidates. The technique can also be applied to the variables in *Sys$'$* where there are no fast variables since the only reaction γ is slow.

Hence the states of the transition systems can be transformed without changing the labels on the transitions. The new transition systems have the same form as the original transition systems, but the new states are vectors with the first three elements of the original vector removed. A new relation can be defined over these new transition systems that preserves the relationship between states. Let $\mathscr{R}' = \big\{ ((k, l, j), (k)) \mid 0 \le k \le n, 0 \le j \le \min\{m, n-k\}, 0 \le l \le p, j+l \le m \big\}$.

Since $\mathscr{R}'$ has the form required for the application of Proposition 6.1, and the two models have the same slow variables, if $\mathscr{R}'$ is a slow bisimulation for $\mathscr{A}_s$, then it is a fast-slow bisimulation for $\mathscr{A}_f$. The new transition system is isomorphic to the original transition system and the relationship between states is preserved by $\mathscr{R}'$, hence $\mathscr{R}$ is also a fast-slow bisimulation for $\mathscr{A}_f$ over the original transition system.

We now proceed with the proof that $\mathscr{R}'$ is a slow bisimulation for $\mathscr{A}_s$. For notational convenience, we let $(P)_i = \{P{:}\uparrow(1, i)\}$, and consider in turn the different cases for which there are γ transitions.

- Consider $((k, l, j), (k)) \in \mathscr{R}'$ for $0 \le k < n$, $0 \le l \le p$, $0 < j \le \min\{m, n-k\}$ which represent states where some substrate-enzyme compound available. Then $(k) \xrightarrow{\gamma, (P)_k} (k+1)$ is matched by $(k, l, j) \xrightarrow{\gamma, (P)_k} (k+1, l, j-1)$ and *vice versa*.

- Consider $((k, l, 0), (k)) \in \mathscr{R}'$ for $0 \le k < n$, $0 \le l \le p$ when no substrate-enzyme compound is present. There are three cases depending on the relationship of m and p.

- If $m > p$, consider $0 \le l \le p$. Since m is greater than p, whatever l is, there will be additional enzyme to form *SE* and $(k) \xrightarrow{\gamma, (P)_k} (k+1)$ is matched by $(k,l,0) \rightarrow (k,l,1) \xrightarrow{\gamma, (P)_k} (k+1,l,0)$.
- If $m \le p$ and $0 \le l \le m-1$, then enzyme is available and the previous case applies.
- If $m \le p$ and $l = m$, then all enzyme is bound in *EI*. Then $(k) \xrightarrow{\gamma, (P)_k} (k+1)$ is matched by $(k,0,m) \twoheadrightarrow (k,0,m-1) \twoheadrightarrow (k,1,m-1) \xrightarrow{\gamma, (P)_k} (k+1,0,m-1)$.

An example of the first subcase is illustrated in Figure 3 in the unmodified transition system. Consider the pair of states $((2,3,0,3,0,0),(2,3,0,3)) \in \mathcal{R}$. The γ-transition from $(2,3,0,3)$ to $(1,3,0,4)$ is matched by a fast transition from $(2,3,0,3,0,0)$ to $(1,2,0,3,0,1)$ and a γ-transition from the latter to $(1,3,0,4,0,0)$ and $((1,3,0,4,0,0),(1,3,0,4)) \in \mathcal{R}$.

To conclude, we have shown for $\{\alpha_1, \alpha_{-1}, \beta_1, \beta_{-1}\} \subseteq \mathcal{A}_f$, $\gamma \in \mathcal{A}_s$ for the models *Sys* and *Sys'* that $(n,m,p,0,0,0) \approx_{\mathcal{A}_f} (n,m,p,0)$ for all positive n, m and p which covers all major cases of interest. Hence, we can conclude that the simpler model demonstrates the same behaviour (at a semi-quantitative level) as the more complex model when we abstract from fast reactions. We can apply the congruence result: if P is a Bio-PEPA model with no fast reactions in $\{\alpha_1, \alpha_{-1}, \beta_1, \beta_{-1}\}$, then since *Sys* $\approx_{\mathcal{A}_f}$ *Sys'*, we know that $P \bowtie_{*} Sys \approx_{\mathcal{A}_f} P \bowtie_{*} Sys'$. This allows us to build new systems, and also to replace the larger state space of *Sys* with the smaller one of *Sys'*.

8 Related and further work

Various approaches to modelling biological systems using process algebra have been proposed including κ-calculus [18], π-calculus [34, 32, 5], Beta-binders [31], Bio-Ambients [33] sCCP [8] and the continuous π-calculus [26]. Most of these approaches use stochastic simulation as their analysis tool, and very few approaches have considered the use of semantic equivalences. Both weak bisimulation and context bisimulation are shown to be congruences for the bio-κ-calculus. Context bisimulation allows for the modelling of cell interaction [27]. Observational equivalence has been used to show that CCS specifications of elements of lactose operon regulation have the same behaviour as more detailed models [30]. In an example of biological modelling using hybrid systems, bisimulation is used to quotient the state space with respect to a subset of variables as a technique for state space reduction [1]. Bisimulation has also been used in the comparison of ambient-style models and membrane-style models [11] and the comparison of a term-rewriting calculus and a simple brane calculus [3]. Other equivalences have been defined for Bio-PEPA. Compression bisimilarity is based on the idea that different discretisations of a system should have the same behaviour assuming sufficient levels [20]. Strong and weak bisimulation parameterised by functions have also been defined [19] and their use demonstrated on a model with alternative pathways. Further work is to determine whether fast-slow bisimilarity can be expressed as g-bisimilarity.

Although fast-slow bisimilarity is defined in the context of Bio-PEPA, it is applicable to any formalism with the same style of stratification of molecular counts or discretisation of concentrations, such as the Petri net-based modelling framework of Heiner *et al* [23].

QSSA has also been applied to stochastic simulation [21] either to obtain approximate rates [9] or in the case of slow-scale stochastic simulation [9, 10] to identity slow and fast species which then leads to the introduction of a virtual fast process representing the fast species where slow reactions are removed.

As mentioned earlier, QSSA is a time scale separation technique. There are other variants such as tQSSA which consider the total substrate (both free and bound) in deriving reduced equations and is applicable when $S_0 + K_m \gg E_T$ does not hold [7]. QSSA approaches have been formalised by single perturbation theory [37, 38].

Another form of time scale decomposition/separation considers CTMCs and is based on a decomposition/ aggregation technique for solving for steady state. In a nearly completely decomposable CTMC, the values in the diagonal blocks are much larger (at least one order of magnitude) than those in the off-diagonal blocks [17]. Hence there are blocks of states where transitions between states in an block is much more frequent than transitions between states in different blocks. The technique involves solving for steady state for each block (ignoring transitions to other blocks). Each block is then considered as a single state, and transition rates between these states are computing, and the aggregate CTMC constructed is solved. Finally, the solutions for each block and the aggregate CTMC are combined to obtain an approximate solution for the original CTMC.

This technique has been applied to both stochastic Petri nets [4] and stochastic process algebra [25]. For Petri nets, a function is defined over markings to determine which markings are similar and must take into account relative rates. In the case of stochastic process algebra, an analysis of processes and the rates of the actions they enable is the starting point for identifying subsets of states. Sequential components are categorised as fast, slow and hybrid, and states are grouped when they have the same slow subcomponents. The passive rate can be used to split hybrid processes into two sequential processes with the same behaviour. A time scale decomposition technique for transient analysis [6] is also relevant because our model considers transient behaviour as well as steady state behaviour. Since we are not working fully quantitatively here, these approaches are issues for further research. Specifically, we wish to compare the application of the technique for nearly completely decomposable CTMCs with a QSSA-based quantitative equivalence, as well as considering transient behaviour.

Quantitative equivalences have been defined for CTMCs based on Kripke structures, hence with unlabelled transitions and labelled states [2]. Both weak bisimulation and weak simulation are defined. Further research involves applying these equivalences, after suitable modification to CTMCs obtained from labelled transition systems and seeing their relationship with the QSSA.

Most previous CTMC research assumes fixed rates; however, with Bio-PEPA rates are state-dependent which introduces additional complexity.

9   Conclusion

We have developed fast-slow bisimilarity, a semi-quantitative semantic equivalence motivated by the Quasi-Steady-State Assumption. We show that for two operators of interest, fast-slow bisimilarity is a congruence. For the cooperation operator, a reasonable condition is required to ensure congruence. The second operator is an extension operator which allows a species to be extended with new reactive capabilities. Although the definition of fast-slow bisimilarity is similar to that of weak bisimilarity, the condition for congruence for cooperation illustrates how they differ. For certains types of reduced models, it is possible to work with slow bisimilarity which only considers slow reactions. The use of fast-slow bisimilarity is illustrated with an example of competitive inhibition, where one system includes the intermediate compounds and the other does not.

This equivalence can be used to show that a reduced system has the same behaviour as the full system. This means it is possible to work only with the reduced system, thereby reducing the number of parameters that need to be fitted. Fast-slow bisimilarity can be applied in any context where concentrations are discretised or molecule counts are grouped.

Further work includes a fully quantitative equivalence, automation of the bisimulation technique including variable reduction and investigation of dynamically changing the sets of fast and slow reactions.

References

[1] M. Antoniotti, C. Piazza, A. Policriti, M. Simeoni & B. Mishra (2004): *Taming the complexity of biochemical models through bisimulation and collapsing: theory and practice*. Theoretical Computer Science 325, pp. 45–67, doi:10.1016/j.tcs.2004.03.064.

[2] C. Baier, J.-P. Katoen, H. Hermanns & V. Wolf (2005): *Comparative branching-time semantics for Markov chains*. Information and Computation 200, pp. 149–214, doi:10.1016/j.ic.2005.03.001.

[3] R. Barbuti, A. Maggiolo-Schettini, P. Milazzo & A. Troina (2008): *Bisimulation in calculi modelling membranes*. Formal Aspects of Computing 20, pp. 351–377, doi:10.1007/s00165-008-0071-x.

[4] A. Blakemore & S.K. Tripathi (1993): *Automated time scale decomposition and analysis of stochastic Petri nets*. In: Proceedings of Petri Nets and Performance Models, pp. 248–257, doi:10.1109/PNPM.1993.393446.

[5] R. Blossey, L. Cardelli & A. Phillips (2006): *A compositional approach to the stochastic dynamics of gene networks*. In: TCSB IV, LNCS 3939, Springer, pp. 99–122, doi:10.1007/11732488_10.

[6] A. Bobbio & K.S. Trivedi (1986): *An Aggregation Technique for the Transient Analysis of Stiff Markov Chains*. IEEE Transaction on Computers 35, pp. 803–814, doi:10.1109/TC.1986.1676840.

[7] J. Borghans, R. de Boer & L. Segel (1996): *Extending the quasi-steady state approximation by changing variables*. Bulletin of Mathematical Biology 58, pp. 43–63, doi:10.1007/BF02458281.

[8] L. Bortolussi & A. Policriti (2008): *Modeling biological systems in stochastic Concurrent Constraint Programming*. Constraints 13, pp. 66–90, doi:10.1007/s10601-007-9034-8.

[9] Y. Cao, D. T. Gillespie & L.R. Petzold (2005): *Accelerated stochastic simulation of the stiff enzyme-substrate reaction*. Journal of Chemical Physics 123(14), p. 144917, doi:10.1063/1.2052596.

[10] Y. Cao, D. T. Gillespie & L.R. Petzold (2005): *The slow-scale stochastic simulation algorithm*. Journal of Chemical Physics 122(1), p. 14116, doi:10.1063/1.1824902.

[11] G. Ciobanu & B. Aman (2008): *On the relationship between membranes and ambients*. BioSystems 91, pp. 515–530, doi:10.1016/j.biosystems.2007.01.006.

[12] F. Ciocchetta & M. L. Guerriero (2009): *Modelling biological compartments in Bio-PEPA*. Electronic Notes in Theoretical Computer Science 227, pp. 77–95, doi:10.1016/j.entcs.2008.12.105.

[13] F. Ciocchetta & J. Hillston (2008): *Process Algebras in Systems Biology*. In: Formal Methods for Computational Systems Biology (SFM08), LNCS 5016, pp. 265–312, doi:10.1007/978-3-540-68894-5_8.

[14] F. Ciocchetta & J. Hillston (2009): *Bio-PEPA: a framework for the modelling and analysis of biological systems*. Theoretical Computer Science, 410(33-34), pp. 3065–3084, doi:10.1016/j.tcs.2009.02.037.

[15] A. Clark, V. Galpin, S. Gilmore, M.L. Guerriero & J. Hillston (2012): *Formal methods for checking the consistency of biological models*. In: Advances in Systems Biology, Advances in Experimental Medicine and Biology 736, Springer.

[16] Allan Clark, Stephen Gilmore, Maria Luisa Guerriero & Peter Kemper (2010): *On verifying Bio-PEPA models*. In: Proceedings of CMSB 2010, ACM Press, pp. 23–32, doi:10.1145/1839764.1839769.

[17] P.J. Courtois (1977): *Decomposability: Queueing and Computer System Applications*. Academic Press.

[18] V. Danos & C. Laneve (2004): *Formal molecular biology*. Theoretical Computer Science 325, pp. 69–110, doi:10.1016/j.tcs.2004.03.065.

[19] V. Galpin (2011): *Equivalences for a biological process algebra*. Theoretical Computer Science, to appear, doi:10.1016/j.tcs.2011.07.006.

[20] V. Galpin & J. Hillston (2011): *A semantic equivalence for Bio-PEPA based on discretisation of continuous values*. Theoretical Computer Science 412, pp. 2142–2161, doi:10.1016/j.tcs.2011.01.007.

[21] D. T. Gillespie (1997): *Exact stochastic simulation of coupled chemical reactions*. Journal of Physical Chemistry 81(25), pp. 2340–2361, doi:10.1021/j100540a008.

[22] C.A. Gómez-Uribe, G.C. Verghese & A.R. Tzafriri (2008): *Enhanced identification and exploitation of time scales for model reduction in stochastic chemical kinetics*. Journal of Chemical Physics 129, doi:10.1063/1.3050350.

[23] M. Heiner, D. Gilbert & R. Donaldson (2008): *Petri nets for Systems and Synthetic Biology*. In: Formal Methods for Computational Systems Biology (SFM08), LNCS 5016, pp. 215–264, doi:10.1007/978-3-540-68894-5_7.

[24] J. Hillston (1996): *A compositional approach to performance modelling*. CUP.

[25] J. Hillston & V. Mertsiotakis (1995): *A Simple Time Scale Decomposition Technique for Stochastic Process Algebras*. Computer Journal 38, pp. 566–577, doi:10.1093/comjnl/38.7.566.

[26] M. Kwiatkowski & I. Stark (2008): *The Continuous π-Calculus: A Process Algebra for Biochemical Modelling*. In: Proceedings of CMSB 2008, LNCS 5307, pp. 103–122, doi:10.1007/978-3-540-88562-7_11.

[27] C. Laneve & F. Tarissan (2008): *A simple calculus for proteins and cells*. Theoretical Computer Science 404, pp. 127–141, doi:10.1016/j.tcs.2008.04.011.

[28] L. Michaelis & M. Menten (1913): *Die Kinetik der Invertinwirkung*. Biochemistry Zeitung 49, pp. 333–369.

[29] R. Milner (1989): *Communication and concurrency*. Prentice Hall.

[30] M.C. Pinto, L. Foss, J.C.M. Mombach & L. Ribeiro (2007): *Modelling, property verification and behavioural equivalence of lactose operon regulation*. Computers in Biology and Medicine 37, pp. 134–148, doi:10.1016/j.compbiomed.2006.01.006.

[31] C. Priami & P. Quaglia (2004): *Beta Binders for Biological Interactions*. In: Proceedings of CMSB 2004, LNCS 3082, pp. 20–33, doi:10.1007/978-3-540-25974-9_3.

[32] C. Priami, A. Regev, E. Shapiro & W. Silverman (2001): *Application of a stochastic name-passing calculus to representation and simulation of molecular processes*. Information Processing Letters 80, pp. 25–31, doi:10.1016/S0020-0190(01)00214-9.

[33] A. Regev, E. Panina, W. Silverman, L. Cardelli & E. Shapiro (2004): *BioAmbients: an abstraction for biological compartments*. Theoretical Computer Science 325, pp. 141–167, doi:10.1016/j.tcs.2004.03.061.

[34] A. Regev & E. Shapiro (2002): *Cellular abstractions: Cells as computation*. Nature 419, p. 343, doi:10.1038/419343a.

[35] K.R. Sanft, D.T. Gillespie & L.R.Petzold (2010): *Legitimacy of the stochastic Michaelis-Menten approximation*. IET Systems Biology 5, pp. 58–69, doi:10.1049/iet-syb.2009.0057.

[36] I. H. Segel (1993): *Enzyme Kinetics: Behaviour and Analysis of Rapid Equilibrium And Steady-State Enzyme Systems*. Wiley Blackwell.

[37] L.A. Segel & M. Slemrod (1989): *The Quasi-Steady-State Assumption: A Case Study in Perturbation*. SIAM Review 31, pp. 446–477, doi:10.1137/1031091.

[38] A. Zagaris, H.G. Kaper & T.J. Kaper (2004): *Analysis of the Computational Singular Perturbation Reduction Method for Chemical Kinetics*. Journal of Nonlinear Science 14, pp. 59–91, doi:10.1007/s00332-003-0582-9.