Reachability in Biochemical Dynamical Systems by Quantitative Discrete Approximation

L. Brim, J. Fabriková, S. Dražan, and D. Šafránek*

Faculty of Informatics
Masaryk University
Botanická 68a, Brno, Czech Republic

safranek@fi.muni.cz

In this paper a novel computational technique for finite discrete approximation of continuous dynamical systems suitable for a significant class of biochemical dynamical systems is introduced. The method is parameterized in order to affect the imposed level of approximation provided that with increasing parameter value the approximation converges to the original continuous system. By employing this approximation technique, we present algorithms solving the reachability problem for biochemical dynamical systems. The presented method and algorithms are evaluated on several exemplary biological models and on a real case study. This is a full version of the paper published in the proceedings of CompMod 2011.

1   Introduction

Under the modern holistic paradigm provided by *systems biology* [5], genome-scale knowledge of individual components is combined with knowledge of interactions underlying the physiology of living organisms. The central goal of systems biology is to integrate all available biological data and to reconstruct *executable models* [19] which allow to investigate complicated behaviour emerging from the underlying biochemistry. An important dimension is the quantitative aspect of the data and processes being modeled.

With respect to [18], we consider biological models to be captured by the notion of a *biochemical dynamical system* consisting of variables describing a certain quantity of the respective species in time (e.g., number of molecules or molar concentration). Variable values evolve in time with respect to rules modeling the effect of reactions. The space of all possible configurations of variable values is referred as the *state space*.

There exist several modeling approaches that differ in abstraction employed for modeling of time, variable values, and molecular interaction effects. The most commonly used approach concerns systems of ordinary differential equations (ODE) [28] where both time and model variables are interpreted as continuous quantities. Effects of interactions are modeled in terms of continuous deterministic updates of variables. Variable values represent molar concentrations of the species. In general, the ODE approach relies on many physical and chemical assumptions simplifying thermodynamic conditions under which particular biochemical phenomena can be modeled correctly [25]. It is important to note that even simple interactions such as second order reactions lead to non-linear ODEs. However, under certain assumptions, biological systems make specific subclasses of general non-linear dynamical systems. Such a specialization motivated development of specific analysis techniques [18, 7, 4, 24].

Nevertheless, dimensionality and complexity of biological models preclude satisfactory application of analysis methods implying that to explore the model dynamics the only practicable method is numerical simulation. Since numerical simulation generates an approximate solution (a trajectory) starting from a single initial point in the continuous state space, the scope of such exploration is limited to the particular trajectory only. This is sufficient for "local" analysis provided that initial conditions are precisely known. However, studied systems are typically under-determined in terms of uncertain quantitative parameters and initial conditions. Therefore generalization of the exploration scope is necessary to reveal and understand the complicated emergent behaviour. An important example of a problem which cannot be effectively solved by local methods is *global temporal property* – the problem to decide whether a given dynamical phenomenon, e.g., oscillation or variables correlation, is globally present/absent for all considered initial conditions [16, 9].

In this paper we limit ourselves to a subclass of dynamical phenomena representing *reachability* of a given portion of the state space. Example of a global temporal property problem that belongs to this subclass is to identify minimal or maximal concentration of species reachable from a particular set of initial conditions.

In general, the reachability problem is undecidable due to unboundedness and uncountability of the state space. However, since concentrations of species cannot expand infinitely, state spaces of biological systems dynamics can be considered bounded in most cases. Analysis can be therefore considered indirectly on suitable finite discrete approximations of continuous state spaces [24, 3].

For a significant class of biochemical dynamical systems determined by multi-affine vector fields (i.e., affine in each variable), there has been developed an over-approximative abstraction technique based on partitioning the continuous state space by a finite set of *rectangles*. Rectangles determine states of a *rectangular transition system* representing the finite discrete (over)approximation of the continuous state space [14], as shown in Figure 1a. The rectangular abstraction has been employed in [24] for reachability analysis and further elaborated by model checking methods in [6]. The results show that the extent of spurious behaviour introduced by the abstraction is typically very high thus limiting satisfactory application of the method. The problem is based mainly on the fact that a transition between any two individual rectangles over-approximates the vector field on the border between the rectangles (a so-called *facet*, see Figure 1b) provided that the information regarding which trajectories starting in an entry facet evolve through a particular exit facet is abstracted out. This causes the rectangular transition system to generate many rectangle sequences in which there is no corresponding trajectory of the original continuous system embedded. Moreover, the extent of such spurious behaviour is not directly eliminated by increasing the partition density.

When analysing approximate models as in systems biology, the need for precise results critically required in systems verification can be relaxed provided that a suitable approximation of the solution can be even more efficient to obtain useful results. Henceforth, in the field of complex systems exhaustive techniques are often combined with approximative methods thus making a certain shift in the way of applying formal methods [29, 11, 10].

1.1   Our Contribution

We present a new technique for discrete approximation of biochemical systems with dynamics given by a system of ODEs with multi-affine right hand side. Our discrete approximation is not an exact abstraction wrt the original continuous system, but rather an approximation that approaches exact reachability with decreasing approximation granularity. While still assuming the rectangular partition at the background, we employ a *measure* that enables local quantification of the amount of trajectories evolving on a rectan-

gle in a particular facet-to-facet direction. To this end, every rectangle is augmented with a local memory representing the information at which part (*entry set*) of the entry facet it has been entered. On each entry set, we identify *focal subsets* from which all trajectories lead to the same exit facet. In Figure 1c, there are two different states of a *quantitative discrete approximation automaton* (QDAA) depicted. Both states share the same rectangle $[1, 1.5] \times [1, 1.5]$ and they differ in entry sets (marked yellow). The upper state with entry set $\{1.5\} \times [1, 1.5]$ has only one focal set - all trajectories from its entry set exit the state through the facet $[1, 1.5] \times \{1\}$. The second state with entry set $[1, 1.5] \times \{1.5\}$ has two focal subsets made by the green and the red part of the entry facet, respectively.

Transitions from a state with given entry set have weights assigned to themselves. Consider a transition from a state A to a state B. The transition exists if there is a part P of the entry set of A such that the trajectories of ODE solutions go from P to B. Weight of a transition from A to B corresponds to the ($n-1$-dimensional) volume of P divided by the volume of the entry set of A. In this manner, the measure reflects amounts of trajectories proceeding in a particular direction. Rectangle regions related by weighted transitions make the QDAA which is a discrete-time Markov chain. (See Theorem 3.2 and its proof.)

From a computational viewpoint, the continuous volumes are finitely approximated by discretization on a uniform grid. Local numerical simulations are employed to identify the entry regions and focal subsets. The density of facet discretization grid is considered as the *method parameter*. Because of combining numerical simulation with rectangular abstraction, the resulting QDAA makes neither an over- nor an under-approximation of the original continuous system. Since for every sequence of states the approximate volume measure converges to the continuous volume with increasing discretization parameter, the parameter indirectly affects the correspondence between the original continuous behaviour and its approximation. This makes the method sufficient for approximating reachability in complex biochemical dynamical systems.

In general, the following main contributions are brought by this paper.

1. A novel computational technique for finite discrete approximation of multi-affine dynamical systems by means of QDAA.

2. Showing that QDAA converges to the original continuous system behaviour. (See Theorem 3.3 and its proof.)

3. A reachability algorithm for QDAA.

4. Evaluation on elementary models and an *E. Coli* case study.

Since the most common application of the considered systems class is the domain of biochemical dynamical systems modeled directly by rules of mass action kinetics [23], evaluation of the method and algorithms is realized on biological models fitting this framework.

1.2 Related Work

Discrete approximation methods are commonly used in continuous and hybrid systems analysis (see [2] for an overview regarding reachability) to handle the uncountability of the state space. Direct methods work on the original system and rely on a successor operation iteratively computing the reachable set whereas indirect methods abstract from the continuous model by a finite structure for which the analysis is simpler. Our method belongs to the latter class, since it uses numerical simulations and creates the abstraction automaton. Considering a fixed set of initial conditions, there is a certain overhead with
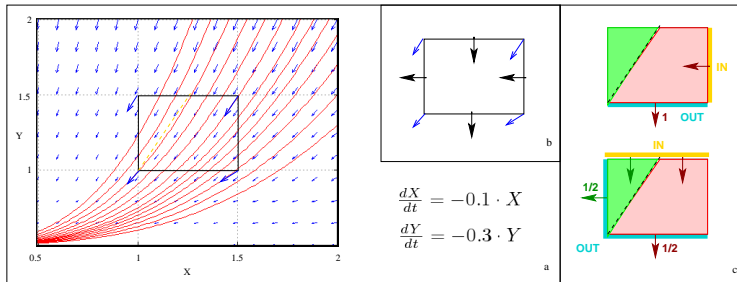
Figure 1: (a) Vector field of a linear system partitioned by thresholds, (b) the principle of rectangular abstraction, (c) and quantification of the extent of over-approximation in terms of transition weights. The dashed line inside the rectangle demonstrates the approximate border separating trajectories exiting through different facets.

generating states of the automaton in comparison with simple numerical simulations. However, the advantage of constructing the automaton is obtaining a global view of the dynamics. Moreover, in addition to rectangular abstraction, the automaton is augmented with weighted transitions which represent quantitative information describing volumes of subsets of initial conditions belonging to attraction basins of different parts of the phase space.

An indirect method based on rectangular abstraction automaton making the finite quotient of the continuous state space has been employed, e.g., in [24, 1, 3]. In general, these methods rely on results [14, 20] and are applicable to (piece-wise) affine or (piece-wise) multi-affine systems. Although not addressed formally in this paper, our technique can be considered as a refinement of [24]. However, we focus on obtaining satisfactory approximate results eliminating the extent of spurious behaviour coming from conservativeness of rectangular abstraction. Our technique can be employed for the recognition of spurious behaviour of the rectangular abstraction transition system.

The technique presented in [27] employs timed automata for the finite quotient of a continuous system as an alternative to piece-wise linear approximations. Another indirect technique adapted to multi-affine biological models is [15]. The approach also employs rectangular abstraction, but results in less conservative reachable sets by means of polyhedral operations. In [2, 8] there are techniques proposed for rectangular refinement that go towards reduction of over-conservativeness. These techniques work fine for linear systems while leaving the non-linear systems as a challenge.

Direct methods are mostly based on hybridization realized by partitioning the system state space into domains where the local continuous behaviour is linearized [12]. This method, in an improved form, has been applied to non-linear biochemical dynamical systems [17]. In general, direct methods give good results for low-dimensional systems and small initial sets. In comparison with indirect approaches, they are computationally harder. From this viewpoint, our approach lies between both extremes.

2 Preliminaries

2.1 Basic definitions and facts

Let $\mathbb{N}$ denote the set of positive integers, $\mathbb{N}_0$ the set $\mathbb{N} \cup \{0\}$, and $\mathbb{R}_0^+$ the set of nonnegative real numbers. For $n \in \mathbb{N}$, denote $\mathbb{R}^n$ the standard n-dimensional Euclidean space with standard topology and Euclidean norm $|\cdot| : \mathbb{R}^n \to \mathbb{R}_0^+$. For an arbitrary function f we use the common notation $dom(f)$ for the domain of

f.

For every $i \in \{1, \ldots, n\}$ assume $a_i, b_i \in \mathbb{R}$ such that $a_i \leq b_i$. Denote $I = \prod_{i=1}^{n}[a_i, b_i]$ an *n-dimensional closed interval* in $\mathbb{R}^n$ and $vol(I)$ the *n-dimensional volume of I* defined as $vol(I) = \prod_{i=1}^{n}(b_i - a_i)$. Further denote *Inter(I)* the *interior of I*, defined as the cartesian product of open intervals $\prod_{i=1}^{n}(a_i, b_i)$.

For any $X \subseteq \mathbb{R}^n$ denote $\lambda_n^*(X)$ the *Lebesgue outer measure* (on $\mathbb{R}^n$) of the set X. Basically $\lambda_n^*(X)$ is the minimal nonnegative real number such that whenever X can be covered by a sequence of closed intervals in $\mathbb{R}^n$ the sum of volumes of these intervals is greater then or equal to $\lambda_n^*(X)$. (For precise definitions see [30].) Note that $\lambda_n^*(X) < \infty$ for every bounded set X and $\lambda_n^*(I) = vol(I)$ for every *n*-dimensional interval I.

Let $n \geq 2, i \leq n, c \in \mathbb{R}$. We use $\mathbb{R}_i^{n-1}(c)$ to denote the hyper-plane $\mathbb{R}_i^{n-1}(c) = \{\langle x_1, \ldots, x_n \rangle \in \mathbb{R}^n \mid x_i = c\}$. Denote $\hat{\pi}_i : \mathbb{R}^n \to \mathbb{R}^{n-1}$ the projection omitting the *i*th variable, $\hat{\pi}_i(\langle x_1, \ldots, x_n \rangle) = \langle x_1, \ldots, x_{i-1}, x_{i+1}, \ldots x_n \rangle$. Let $X \subseteq \mathbb{R}_i^{n-1}(c)$. We extend the notion of the $(n-1)$-dimensional Lebesgue outer measure to such sets X and denote $\lambda_{n-1}^*(X)$ the $(n-1)$-dimensional Lebesgue outer measure of $\hat{\pi}_i(X)$.

Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous function (an autonomous vector field). We say that

$$\dot{x} = f(x) \tag{1}$$

is an *autonomous ODE system*. An important property of autonomous systems is the fact that if $y(t)$ is a solution of (1) on an open interval (a, b), then $y(t + t_0)$ is also a solution (defined on interval $(a - t_0, b - t_0)$).

A function $f : \mathbb{R}^n \to \mathbb{R}^n$ *satisfies the Lipschitz condition locally* on $\mathbb{R}^n$, if for every $x \in \mathbb{R}^n$ there exists an open set $U \subseteq \mathbb{R}^n$, $x \in U$ and a constant $L \in \mathbb{R}$ such that for every two points $x_1, x_2 \in U$ the inequality $|f(x_1) - f(x_2)| \leq L \cdot |x_1 - x_2|$ holds.

Theorem 2.1 (Trajectories of solutions of an autonomous system) *Let (1) be an autonomous ODE system, where f is defined on $\mathbb{R}^n$ and let f satisfy the Lipschitz condition locally on $\mathbb{R}^n$. Let x be an inextendible solution of system (1). Then $dom(x)$ is an open interval, and for every point $\alpha \in \mathbb{R}^n$ there exists exactly one trajectory of an inextendible solution $x(t)$ of system (1) coming through α.*

Theorem 2.2 (Continuous dependency on initial conditions) *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be continuous on an open set $E \subseteq \mathbb{R}^n$ with the property that for every $y_0 \in E$, the initial value problem $\dot{x} = f(x), x(0) = y_0$ has a unique solution $y(t) = \eta(t, y_0)$ (η is a function of variables t, y_0). Let $w_\perp, w_\top \in \mathbb{R}$ such that $(w_\perp, w_\top)$ is the maximal interval of existence of $y(t) = \eta(t, y_0)$.*

Then the bounds $w_\perp, w_\top$ are (lower, resp. upper semicontinuous) functions of y_0 in E and $\eta(t, y_0)$ is continuous on the set $\{\langle t, y_0 \rangle \mid y_0 \in E, w_\perp(y_0) < t < w_\top(y_0)\} \subseteq \mathbb{R}^{n+1}$.

We restrict ourselves to multi-affine autonomous systems. That is, systems of the form (1), such that the vector field f is a *multi-affine* function, defined as a polynomial of variables $x_1, \ldots, x_n \in \mathbb{R}^n$ of degree at most one in every variable. The assumptions of Theorems 2.1 and 2.2 (from [21]) are satisfied for systems of this class, therefore the properties stated in the above theorems can be used for reasoning about autonomous systems with multi-affine vector fields.

2.2 Biochemical dynamical system

According to [18], by a biochemical dynamical system we understand a collection of n biochemical species interacting in biochemical reactions. Species concentrations are represented by variables $x_1, \ldots, x_n$ attaining values from $\mathbb{R}_0^+$. If the stoichiometric coefficients in reactions do not exceed one

and the reaction dynamics respects the law of mass action kinetics [23], the dynamical system can be described by a multi-affine autonomous system in the form (1).

In a biochemical dynamical system we are typically interested in a bounded part (n-dimensional interval) of the phase space in $\mathbb{R}^n$. Further, we consider the phase space partitioned by a (non-uniform) rectangular grid. In particular, for each variable there is defined a finite set of *thresholds*, making the system *partition*. Thresholds determine $(n-1)$-dimensional hyper-planes in $\mathbb{R}^n$ and can be freely specified according to particular questions that should be addressed by the model analysis, e.g., specification of unsafe or attracting sets. Cells laid out by $2n$ adjacent threshold hyper-planes (cells are again intervals in $\mathbb{R}^n$) are called *hyper-rectangles*, for short we refer to them as *rectangles*.

Definition 2.1 *Define a* biochemical dynamical system *(*biochemical system *for short) as a tuple* $\mathscr{B} = \langle n, f, \mathscr{T}, \mathscr{I}_C \rangle$*, where*

- *$n \in \mathbb{N}$ is the* dimension *of $\mathscr{B}$,*

- *$f : \mathbb{R}^n \to \mathbb{R}^n$ is the multi-affine vector field of $\mathscr{B}$,*

- *$\mathscr{T} = \langle T_1, \ldots, T_n \rangle$ is the* partition *of $\mathscr{B}$ where each T_i is a finite subset of $\mathbb{R}_0^+$, and define the* set of rectangles *given by $\mathscr{T}$ as*

$$Rect(\mathscr{T}) = \{\prod_{j=1}^n I_j \mid \forall j \exists a, b \in T_j : I_j = [a, b], \forall c \in T_j : c \leq a \lor c \geq b\},$$

- *$\mathscr{I}_C \subseteq Rect(\mathscr{T})$ is the set of* initial conditions *(*initial set*) of $\mathscr{B}$.*

Definition 2.2 *Let $\mathscr{B} = \langle n, f, \mathscr{T}, \mathscr{I}_C \rangle$ be a biochemical system and let $H \in Rect(\mathscr{T})$ be a rectangle such that $H = I_1 \times \ldots \times I_n$, where $I_i = [a_i, b_i]$. For every $i \in \{1, \ldots, n\}$ define the* lower *(resp.* upper*) facet of H wrt the ith variable:*

$$Facet_i^\perp(H) = \{\langle x_1, \ldots, x_n \rangle \in H \mid x_i = a_i\},$$
$$Facet_i^\top(H) = \{\langle x_1, \ldots, x_n \rangle \in H \mid x_i = b_i\}.$$

Denote $Facets_i(H)$ the set of ith dimension facets of H, $Facets_i(H) = Facet_i^\perp(H) \cup Facet_i^\top(H)$, and $Facets(H)$ the set of (all) facets of H, $Facets(H) = \bigcup_{i=1}^n Facets_i(H)$.*

Definition 2.3 *Let $H, H' \in Rect(\mathscr{T})$. We say that H is a* neighbour *of H', denoted $H \bowtie H'$, if there exists $F \in \text{Facets}(H)$ such that $H \cap H' = F$.*

3 Quantitative Discrete Approximation

Given a biochemical system $\mathscr{B} = \langle n, f, \mathscr{T}, \mathscr{I} \rangle$, we aim to define a finite automaton reflecting the behaviour of $\mathscr{B}$, and for each state, to assign every transition a weight quantifying probability of proceeding to a particular successor.

A state is defined as a pair $\langle H, E \rangle$ – a rectangle H, and a subset E of a particular facet of H. The set E represents a so-called *entry set*, a region through which trajectories of the system (1) enter the interior of H. Intuitively, we can say that E encodes the history of previous evolution of the system from initial set $\mathscr{I}_C$ to H. Entry sets are either subsets of $(n-1)$-dimensional facets of H or (in case of initial states) the whole n-dimensional rectangle H.

Since entry sets can be arbitrary sets in Euclidean space, we approximate them by a finite discrete structure. Each facet is provided with a uniform grid on which we approximate any subset of the facet

$$\frac{d[A]}{dt} = -0.5 \cdot [A] + 0.8 \cdot [B]$$

$A \xrightarrow{0.5} B$

$B \xrightarrow{0.8} A$

$$\frac{d[B]}{dt} = 0.5 \cdot [A] - 0.8 \cdot [B]$$

thresholds on $[A]$: $\{0, 2.5, 5\}$
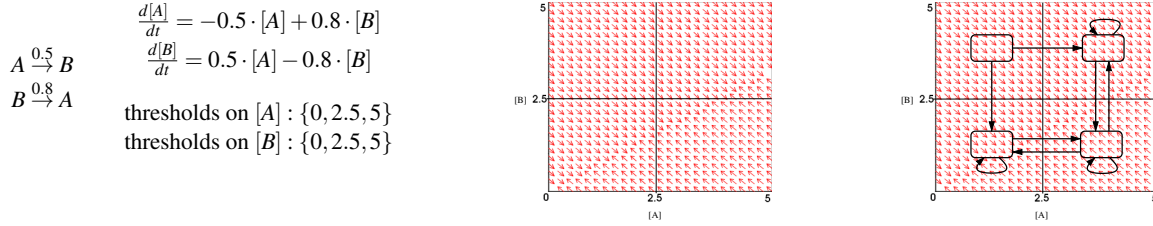thresholds on $[B]$: $\{0, 2.5, 5\}$

Figure 2: Example of a biochemical system with two species and two reactions. Dynamics given by a system of two ODEs and the system of thresholds are in the left part of the figure. Vector field is visualized in the middle, and its Rectangular Abstraction Transition System on the right.

by the set of rectangular fragments, so-called *tiles* (Figure 3). The grid is *n*-dimensional or $(n-1)$ dimensional depending on the dimension of approximated entry sets. When following the trajectories of solutions of differential equations of the models dynamics in time, entry sets are identified by trajectories of solutions passing through them on their way from preceding rectangles. In following definitions we treat this intuitive perception of entry sets formally.

Let $\kappa \in \mathbb{N}$, let $\mathscr{B} = \langle n, f, \mathscr{T}, \mathscr{I}_C \rangle$ be a biochemical system, $H \in Rect(\mathscr{T})$, and $F \in Facets(H)$ for all definitions and theorems from this section.

Definition 3.1 *Let H be of the form $H = \prod_{j=1}^{n} I_j$, where $\forall j : I_j = [a_j, b_j]$. Let $B \in \{H\} \cup Facets(H)$. Set either $n' = n$, if $B = H$, or $n' = (n-1)$, if $B \in Facets_i(H)$ for some $1 \leq i \leq n$ (in this case $\exists c \in \{a_i, b_i\} : B \subset \mathbb{R}_i^{n-1}(c)$).*

Define the set of κ-tiles of B as $Tiles_{n'}^{\kappa}(B) = \{A \subseteq B \mid A = \prod_{j=1}^{n} A_j\}$, where $A_i = \{c\}$, if $B \in Facets_i(H)$, and otherwise ($j \neq i$ or $B = H$) A_j is a closed interval in $\mathbb{R}_0^{+}$ of the form $[a_j + \frac{k_j}{\kappa}(b_j - a_j), a_j + \frac{k_j+1}{\kappa}(b_j - a_j)]$, where for all $j \in \{1, \ldots, n\}, j \neq i$ the nonnegative integer $k_j \in \mathbb{N}_0$ satisfies $k_j < \kappa$.

The following definition introduces the notion of general entry sets.

Definition 3.2 *Define the set of entry points into a rectangle H through facet F, as the set $Entry(F,H) = \{y_0 \in F \mid \exists$ a trajectory $y(t)$ of a solution of (1) such that $y(0) = y_0$ and $\exists \varepsilon > 0 : y(t) \in H$ for $\forall t \in (0, \varepsilon)\}$.*

Next we define the approximation of entry sets on a grid of κ-tiles. Additionally, we define the respective (discrete) volume measure of a set (see Figure 3 c),d)).

Definition 3.3 *Let $X \subset H$. Let $n' = n - 1$, if there exists $i \in \{1, \ldots, n\}, F \in Facets_i(H)$ such that $X \subseteq F$, and let $n' = n$, otherwise. Let $M = F$, if $X \subseteq F$, and let $M = H$, if there is no such facet F. Define*

- *the set of κ-tiles approximating the set X as*

$$Tiles_{n'}^{\kappa}(X) = \left\{ A \in Tiles_{n'}^{\kappa}(M) \mid \frac{\lambda_{n'}^{*}(A \cap X)}{\lambda_{n'}^{*}(A)} \geq \frac{1}{2} \right\},$$

- *the rectangular κ-grid measure of the set X as $\lambda_{n'}^{\kappa}(X) = \sum_{A \in Tiles_{n'}^{\kappa}(X)} vol(A)$.*

The following definition declares the set of all discretized entry sets for a given rectangle.

a)   b)   c)   d)
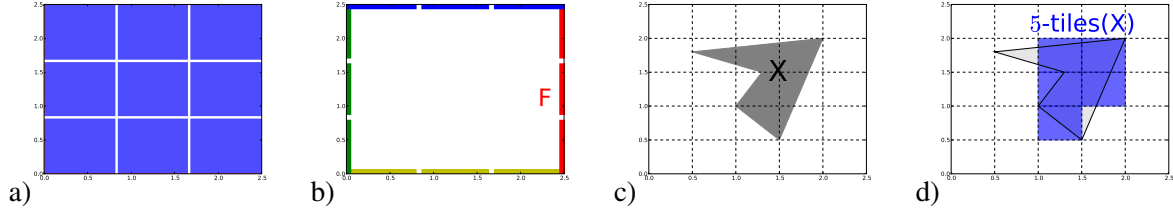
Figure 3: a) Let $H = [0, 2.5] \times [0, 2.5]$ be a rectangle. The blue areas depict elements of $Tiles_2^3(H)$.
b) Let $F = Facet_1^\top(H) = \{2.5\} \times [0, 2.5]$. The red line segments are elements of $Tiles_1^3(F)$.
The set $EntrySets_3(H)$ has $2 + 4 \cdot (1 + \binom{3}{2} + \binom{3}{1}) = 30$ elements: $\emptyset, H$, and 7 for every facet of H (the facet itself, 3 segments and 3 unions of pairs of segments of the facet).
c) Let X be a subset of H (the shaded polygon). Let $\kappa = 5$.
d) The set of κ−tiles approximating X is the set of five blue intervals (each satisfying that at least half of its area is in X). The cardinality of $Tiles_2^\kappa(X)$ is 5. Thus $\lambda_2^\kappa(X) = 5 \cdot (0.5 \cdot 0.5) = 1.25$.

Definition 3.4 *For H, define* set of (approximate) entry sets $EntrySets_\kappa(H) =$

$$\left\{ E \subseteq H \mid E = \emptyset \vee E = H \vee \exists F \in Facets(H), \mathcal{E} \subseteq Tiles_\kappa\big(Entry(F, H)\big) : E = \bigcup \mathcal{E} \right\}.$$

For an example of a set of (approximate) entry sets of a rectangle see Figure 3 a),b). Note that set of approximate entry sets is always finite. Further note that also the empty set and the entire rectangle are considered as entry sets. These represent singular cases needed in the subsequent construction of the automaton. In particular, states with the empty entry set approximate fixed point behaviour not leaving the rectangle (steady state memory) whereas the rectangle-form entry set is employed for initial rectangles.

Definition 3.5 *Let $E \in EntrySets_\kappa(H), H' \in Rect(\mathcal{T}), F' \in Facets(H)$ such that $H' \bowtie H, F' = H \cap H'$.*
Define the focal subset of E on H targeting F', *denoted $Focal(H, E, F')$, as the set of all $y_0 \in E$ such that there exist $\varepsilon, \varepsilon', c > 0$ and a trajectory of a solution $y(t)$ of system (1) with inital conditions $y(0) = y_0$ satisfying $y(t) \in H$ for $t \in (0, c), y(t) \in Inter(H)$ for $t \in (0, \varepsilon)$, $y(c) \in F'$, and $y(t) \in Inter(H')$ for $t \in (c, c + \varepsilon')$. Let $ExitSet(H, E, F')$ denote the set of all such (targeted) points $y(c) \in F'$.*
Define focal subset of E on H not leaving H, $Focal(H, E, \emptyset)$, *as the set of all points $y_0 \in E$ such that there exists a trajectory of a solution $y(t)$ of system (1) with initial conditions $y(0) = y_0$ satisfying $y(t) \in H$ for all $t > 0$.*

Next we define the successor function for any pair $\langle H, E \rangle$ and subsequently the quantitative discrete approximation automaton.

Definition 3.6 *Let $E \in EntrySets_\kappa(H)$. Define the* successors of $\langle H, E \rangle$ *as the set of pairs $\langle H', E' \rangle$ with $H' \in Rect(\mathcal{T}), E' \in EntrySets_\kappa(H')$ such that*

$$Succs(\langle H, E \rangle) = \left\{ \langle H', E' \rangle \mid H', E' \text{ satisfy one of conditions } 1. - 3. \text{ below} \right\}$$

1. $H' \bowtie H, E \neq \emptyset$. Denote F' the facet of H satisfying $F' = H \cap H'$. Let $n' = n$, if $E = H$, and $n' = (n - 1)$, otherwise. Moreover, $E' = \bigcup Tiles_\kappa\big(ExitSet(H, E, F')\big)$ and $\lambda_{n'}^\kappa\big(Focal(H, E, \emptyset)\big) > 0$.
2. $H' = H, E \neq \emptyset$, and $E' = \emptyset$. Further, it holds that either $E \subseteq F$ and $\lambda_{n-1}^\kappa\big(Focal(H, E, \emptyset)\big) > 0$, or $E = H$ and $\lambda_n^\kappa\big(Focal(H, E, \emptyset)\big) > 0$.
3. $H' = H$ and $E' = E = \emptyset$.

Definition 3.7 (The Quantitative Discrete Approximation Automaton) *Let $\kappa, \mathcal{B}$ be as above. The quantitative abstraction automaton $QDAA_\kappa(\mathcal{B})$ of a biochemical system $\mathcal{B}$ with parameter κ is a tuple $QDAA_\kappa(\mathcal{B}) = \langle S, \mathscr{I}_C, \delta, p \rangle$, where*

- *the* set of states $S = \{\langle H, E \rangle \mid H \in Rect(\mathscr{T}), E \in EntrySets_\kappa(H)\}$,
- *the* set of initial conditions $I_C = \{\langle H, H \rangle \mid H \in \mathscr{I}_C\}$,
- *the* transition function $\delta : S \to 2^S$ *is defined as* $\delta(\langle H, E \rangle) = Succs(\langle H, E \rangle)$,
- *the* weight function $p : S \times S \to [0,1]$ *is defined by the following expression, where $S = \langle H, E \rangle, S' = \langle H', E' \rangle$. Suppose $n' = n$, in case $E = H$, and $n' = n - 1$, otherwise.*

$$p(S,S') = \begin{cases} 1, & \text{if } H = H', E = E' = \emptyset, \\[2mm] \dfrac{\lambda_{n'}^*\big(Focal(H,E,\emptyset)\big)}{\sum_{A \in Facets(H) \cup \{\emptyset\}} \lambda_{n'}^*\big(Focal(H,E,A)\big)}, & \text{if } H = H', E \neq \emptyset, E' = \emptyset, \\[4mm] \dfrac{\lambda_{n'}^*\big(Focal(H,E,F')\big)}{\sum_{A \in Facets(H) \cup \{\emptyset\}} \lambda_{n'}^*\big(Focal(H,E,A)\big)}, & \text{if } H \bowtie H', E' \subseteq F' = H \cap H', \\[4mm] 0, & \text{otherwise.} \end{cases}$$

Example 3.1 *Assume the biochemical system from Figure 2. See Figure 4 a) for an example of focal subsets described below. Let $R = [0,2.5] \times [2.5,5]$ be a rectangle and let $F_0 = Facet_2^\top(R), F_1 = Facet_1^\top(R), F_2 = Facet_2^\perp(R), F_3 = Facet_1^\perp(R)$. For the state $\langle R, F_0 \rangle$ the focal set of F_1 equals F_0, whereas $Focal(F_0) = Focal(F_2) = Focal(F_3) = \emptyset$.*

Let $H = [0,2.5] \times [0,2.5]$ and $F = Facet_1^\top(R)$. For the state $\langle H, H \rangle$ the set $Focal(F)$ is the blue area inside H and $Focal(\emptyset)$ is the yellow area. All the solutions of the biochemical systems dynamics with initial conditions in $Focal(\emptyset)$ approach the yellow line of fixed points and stay in H forever. All the solutions starting in the blue area leave H in finite time through F.

In the right part of Figure 4 is the set of reachable states of the quantitative discrete approximation automaton (QDAA) obtained from the biochemical system described in Figure 2 with initial conditions $\mathscr{I}_C = \{[0,2.5] \times [0,2.5]\}$.

Let H, R be the same as above. Let $S = [2.5,5] \times [0,2.5]$ and let $\mathscr{I}_C = \{H\}$. The QDAA successor states of $\langle H, H \rangle$ are $\langle H, \emptyset \rangle$ (a selfloop state) and $\langle S, E \rangle$ (where E denotes the κ-tiles approximation of the red segment in $Facet_1^\perp(S)$). For $\kappa \to \infty$ the weights of these two transitions approach the area ratios of yellow and blue regions of H respectively. The only successor of $\langle H, \emptyset \rangle$ is (by definition) itself. The state $\langle S, E \rangle$ has one successor $\langle S, \emptyset \rangle$, since all the trajectories beginning in E approach the line of fixed points and stay inside S forever.

Therefore the set of concentrations reachable from initial rectangle H is $[0,5] \times [0,2.5]$. See the rectangular abstraction transition system from Figure 2 where the set reachable from H is $[0,5] \times [0,2.5] \cup [2.5,5] \times [2.5,5]$, although there exists no trajectory of a solution of the biochemical systems dynamics that starts in H and reaches a point inside $[2.5,5] \times [2.5,5]$.

On the other hand, if κ is too small, some behaviours of the system are not reflected in QDAA, because the set of κ-tiles corresponding to the entry set may be empty. With finer partition into κ-tiles smaller entry sets can be captured and approximation of the biochemical system by a QDAA is more realistic.

In the Theorem 3.1 we ensure correctness of using the Lebesque measure in Definition 3.7. We ensure that there is no non-zero volume entry set such that all trajectories from this set lead to a facet without entering the interior of a neighbouring rectangle.

Following notations and lemmas provide the background that is used in the proof of Theorem 3.1.
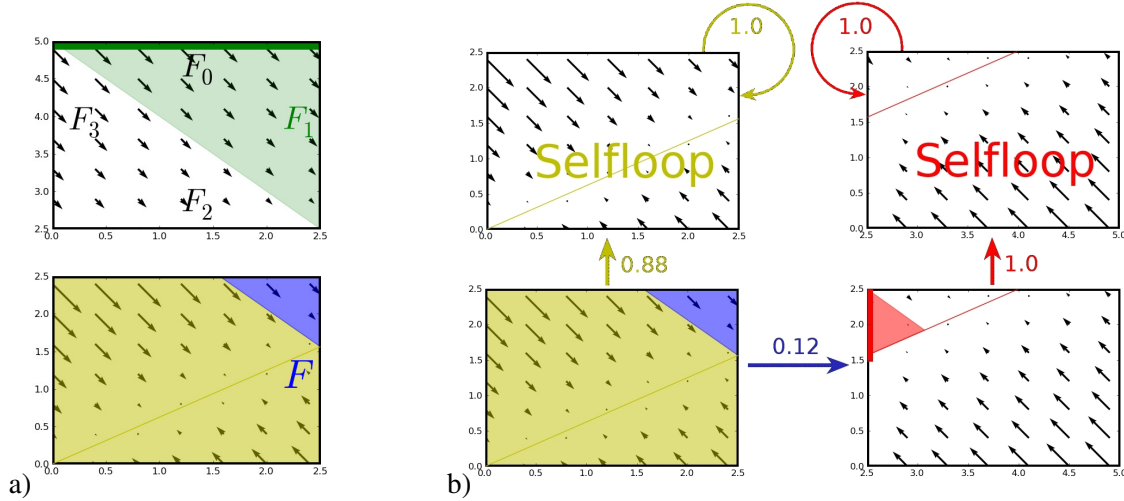
Figure 4: a) Focal sets examples, b) QDAA example.

Notation 3.1 *Let us denote $B_\varepsilon(x)$ the* open sphere *with centre $x \in \mathbb{R}^n$ and radius $\varepsilon > 0$ ($B_\varepsilon(x) = \{y \in \mathbb{R}^n \,|\, |x - y| < \varepsilon\}$).*

The following theorems from mathematical analysis and differential topology (see for example [22], [21]) will be used in the proofs.

Lemma 3.1 (Peano) *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be continuous on open set $E \subseteq \mathbb{R}^n$ such that f possesses continuous first order partial derivatives. Then for every $y_0 \in E$, the initial value problem $\dot{x} = f(x), x(0) = y_0$ has a unique solution $y(t) = \eta(t, y_0)$ and the unique solution has continuous first order derivative wrt the variable t (is of class C^1) on its open domain of definition.*

Lemma 3.2 *The following statements about zero measure sets hold. (A set has Lebesgue outer measure zero iff it has Lebesgue measure zero.)*

1. *If $S \subset \mathbb{R}^n, \lambda_n^*(S) = 0$ and $f : \mathbb{R}^n \to \mathbb{R}^n$ is a smooth map (with continuous first order derivatives), then $\lambda_n^*(f(S)) = 0$.*
2. *Being of zero measure is a diffeomorphism-invariant property of subsets of $\mathbb{R}^n$.*

Lemma 3.3 (A version of Fubini theorem) *Let U be a compact subset of $\mathbb{R}^n$. Denote by $\mathbb{R}_t^n$ the subset $\{t\} \times \mathbb{R}^{n-1} \subset \mathbb{R}^n$, and denote $U_t = U \cap \mathbb{R}_t^n$. If every set U_t satisfies $\lambda_{n-1}^*(U_t) = 0$, then $\lambda_n^*(U) = 0$.*

Lemma 3.4 *Let $g : \mathbb{R}^n \to \mathbb{R}^n$ be a multiaffine function. Let $F \in Facets_i(H)$. The following statements hold.*

1. *If there exists a set $U \subseteq H$ such that $\lambda_n^*(U) > 0$ and $\forall x \in U : g(x) = 0$, then $g \equiv 0$ on H (and on $\mathbb{R}^n$).*
2. *If there exists a set $V \subseteq F$ such that $\lambda_{n-1}^*(V) > 0$ and $\forall x \in V : g(x) = 0$, then $g \equiv 0$ on F (and on the hypherplane $\mathbb{R}_i^n(c)$ containing F).*

Proof The proof of statement *1.* will be done by mathematical induction wrt n.

For $n = 1$ is H a line segment, and the function g is a linear function of one variable. Either there is one point (a set of measure zero in $\mathbb{R}^1$) where the function attains zero value, or the function is zero on the whole line.

For $n = 2$, H is a rectangle in plane. The multi-affine function of two variables is either zero on whole plane, or on two intersecting lines, or a hyperbolic curve or empty set. Therefore either the function is zero on the whole plane, or on a subset of plane of measure zero. Only the second case is compatible with g being nonzero on a set of non-zero Lebesgue outer measure. Assume that the statement holds for $1, 2, \ldots, n-1$ and let us prove it for n. By Theorem 3.3.

Proof of 2. is analogous, we can identify $\mathbb{R}_i^n(c)$ with $\mathbb{R}^{n-1}$ and assume that g is a multiaffine function of $n-1$ variables (with $x_i = c$ constant). $\square$

Definition 3.8 *Let $U \subset H$, let I be an (bounded or unbounded) interval in $\mathbb{R}$. Define the I-trajectories* unoin *of U as the set*

$$\Xi_I(U)\big\{x(s)|s \in I, x(t) \text{ is a solution of system (1) defined on } I, x(0) \in U\big\}.$$

Lemma 3.5 *Let $U \subset H$ be a $(n-1)$-dimensional closed disc in H (i.e.*

$$\exists i \in \{1, \ldots, n\} \exists c \in \mathbb{R} \exists \delta > 0 \exists x \in H : U = \{y \in \mathbb{R}_i^n(c)| |x - y| \leq \delta\},$$

we do not assume there is a facet F of H with $U \subseteq F$).

Let $\lambda_{n-1}^*(U) > 0$ and let either $\forall y \in U : f(y) \cdot e_i > 0$ or $\forall y \in U : f(y) \cdot e_i < 0$.

Then either $\exists U' \subseteq U$ with $\lambda_{n-1}^*(U') > 0$ such that all the solutions of system (1) stay forever in H, or there exists $F' \in Facets(H)$ such that $\lambda_{n-1}^*(F' \cap \Xi_{[0,+\infty)}(U)) > 0$.

Moreover, in the second case, $\lambda_{n-1}^*(F' \cap W) > 0$, where W is the set of points of the boundary of H where the trajectories leave the rectangle H for the first time (a subset of the connected component of $\Xi_{[0,+\infty)}(U) \cap H$ containing U).

Proof Consider the case, when there is no $U' \subseteq U$ with $\lambda_{n-1}^*(U') > 0$ such that all the solutions of system (1) stay forever in H.

Since all compactly supported vector fields are complete, we can restrict ourselves on the multi-affine vector field f defined on a compact neighbourhood of H and assume that flow of this (complete) vector field defines a one-parametric group of diffeomorphisms $Fl_t^f : (t, y_0) \to y(t)$, where $y(t)$ is the respective solution of system 1.

We consider a diffeomorphism φ_U between $\Xi_{[0,+\infty)}(U) \cap H$ and a special subset V of $U \times \mathbb{R}$. We prove that for U satisfying the assumptions of Lemma 3.5 the set V is of nonzero measure. Therefore its image by diffeomorphism φ must be of nonzero measure (due to Theorem 3.2). If on the other hand the set $\Xi_{(-\infty,0]}(F' \cap W) \cap H$ was of measure zero, its image in $\varphi_{F' \cap W}$ must be of zero measure also, but since $\Xi_{[0,+\infty)}(U) \cap H$ is the union of finitely many such sets $\Xi_{(-\infty,0]}(F' \cap W) \cap H$, there has to exist F' such that the set $\Xi_{(-\infty,0]}(F' \cap W) \cap H$ has nonzero measure. $\square$

Lemma 3.6 *Let $E \in EntrySets_\kappa(F, H), E \neq \emptyset$.*

Then either $\exists U' \subseteq E$ with $\lambda_{n-1}^*(U') > 0$ such that all the solutions of system (1) stay forever in H, or there exists $F' \in Facets(H)$ such that

$$\lambda_{n-1}^*(F' \cap \Xi_{[0,+\infty)}(E \cap Entry(F, H))) > 0.$$

Moreover, in the second case, $\lambda_{n-1}^*(F' \cap W) > 0$, where W is the set of points of the boundary of H where the trajectories leave the rectangle H for the first time (a subset of the connected component of $\Xi_{[0,+\infty)}(E \cap Entry(F, H)) \cap H$ containing the set $E \cap Entry(F, H)$).

Proof Consider again the case, when there is no $U' \subseteq U$ with $\lambda^*_{n-1}(U') > 0$ such that all the solutions of system (1) stay forever in H.

The situation is easier if there exists a point $x \in E$ with $f(x) \cdot v_H(F) < 0$. The multi-affine function f is continuous, therefore there exists am open sphere containing x such that $f(y) \cdot v_H(F) < 0$ for all points y from this sphere. As a subset of the intersection of this sphere and F there exists a closed disc like U from the assumption of Lemma 3.5 and the statement of this lemma holds.

Otherwise the whole set $E \cap Entry(F, H)$ is a subset of F with nonzero $(n-1)$-dimensional outer Lebesgue measure on which $f \equiv 0$. By Lemma 3.4 $f \equiv 0$ on the whole $(n-1)$-dimensional hypherplane containing F.

The idea of proof for this configuration is to use the definitorial property of $Entry(F, H)$, that insures existence of ε_z and a point $y \in Inter(H)$ reachable by a trajectory from a point z in $E \cap Entry(F, H)$ in time ε_z with $f(z) \cdot v_H(F) < 0$ (that is implied by the properties of multi-affine function f). By continuous dependency on the initial conditions we find a disc around z that cosists of points reachable from $E \cap Entry(F, H)$ (continuous dependence) and satisfies assumptions of Lemma 3.5. Then we use Lemma 3.5 and this lemma is proved. $\square$

Theorem 3.1 *Let $E \in EntrySets_\kappa(H), E \neq \emptyset$. Further, let $n' = n$, if $E = H$, and $n' = (n-1)$, otherwise. Then*

$$\sum_{A \in Facets(H) \cup \{\emptyset\}} \lambda^*_{n'}\big(Focal(H, E, A)\big) > 0, \tag{2}$$

Proof The proof will be divided into two parts.

First, for $E \subseteq F, E \neq \emptyset$. Second, for $E = H$.

Proof of 1. The statement of theorem is obtained by using Lemma 3.6.

Proof of 2. Let $E = H$. In the case when all the trajectories with initial points in H stay forever in H the inequality $\lambda^*_n\big(Focal(H, E, \emptyset)\big) > 0$ holds.

In the case when there exists a point $x_0 \in H$ and a trajectory of a solution $x(t)$ of system (1) with $x(0) = x_0, \exists c, \varepsilon > 0 : \forall t \in [0, c], x(t) \in H, \forall t \in (c, c + \varepsilon) x(t) \notin H$. Let us denote y the point $x(c + \frac{\varepsilon}{2}) \notin H$. From Theorem 2.2 (continuous dependency on initial conditions) there exists an n-dimensional sphere B with centre x_0 such that trajectories of all solutions of (1) with initial point in the sphere B leave the rectangle H and continue into a neighbourhood of point y. The intersection $H \cap B$ surely contains a disc satisfying assumptions of Lemma 3.5 and therefore there exists $F' \in Facets(H)$ such that the intersection of $\Xi_{[0, +\infty)(B)}$ and F' is of nonzero measure, i.e. it has a subset of nonzero $(n-1)$-dimensional measure in the interior of the facet F'. Then there exists an open sphere B' contained in B such that all trajectories of solutions with initial points in B' leave the rectangle H through the interior of facet F', and $\lambda^*_n\big(Focal(H, E, F')\big) \geq \lambda^*_n(B') > 0$. $\square$

Theorem 3.2 *The quantitative abstraction automaton $QDAA_\kappa(\mathscr{B})$ of a biochemical system $\mathscr{B}$ is a discrete time Markov chain.*

Proof The number of states is finite, bounded by $|Rect(\mathscr{T})| \Big(2 + 2n \cdot \big(2^{\kappa^{n-1}} - 1\big)\Big)$.

Sum of probabilities of transitions from one state $\langle H, E \rangle$ equals 1 for $E = \emptyset$ and $\frac{\sum_{A \in Facets(H) \cup \{\emptyset\}} \lambda^*_{n'}\big(Focal(H, E, A)\big)}{\sum_{A \in Facets(H) \cup \{\emptyset\}} \lambda^*_{n'}\big(Focal(H, E, A)\big)}$. The later sum equals 1, whenever its denominator is nonzero (it is the case because of Theorem 3.1). The probabilities of transitions from a given state are independent of previous states of the automaton. $\square$

Finally, we provide a theorem suggesting that for sufficiently large values of parameter κ, the rectangular κ-grid measure of a bounded set X contained in the phase space of biochemical system approaches

its Lebesque outer measure (Theorem 3.3). Following lemmas and definition will be used in the proof of this theorem.

Let $\mathscr{B} = \langle n, f, T, I_C \rangle$ be a biochemical system, $H \in Rect(\mathscr{T})$, $F \in Facets(H)$, and $\kappa \in \mathbb{N}$ throughout this section.

Definition 3.9 *Let* $X \subseteq \prod_{i=1}^{n}[\min(T_i), \max(T_i)]$. *Define* $\lambda_n^{\kappa}(X)$ *as the sum*

$$\sum_{H \in Rect(\mathscr{T})} \lambda_n^{\kappa}(X \cap H).$$

Analogously define $\lambda_{n-1}^{\kappa}(X)$ *as the sum*

$$\sum_{F \in \bigcup_{H \in Rect(\mathscr{T})} Facets(H)} \lambda_{n-1}^{\kappa}(X \cap F).$$

Lemma 3.7 *Let* $m \in \{n-1, n\}$, *let J be an m-dimensional interval in n-dimensional space,* $J \subseteq \prod_{i=1}^{n}[\min(T_i), \max(T_i)]$. *Then*

$$\lim_{\kappa \to \infty} \lambda_m^{\kappa}(J) = \lambda_m^{*}(J)$$

Proof For given κ there are at most $2n \cdot \kappa^{n-1}$ distinct tiles R satisfying $R \nsubseteq J$ and $R \cap J \neq \emptyset$.

The difference between $\lambda_m^{\kappa}(J)$ and $\lambda_m^{*}(J) = vol(J)$ is bounded by

$$\frac{1}{2} \sum vol(R_i) = \frac{1}{2} 2n \cdot \kappa^{n-1} vol(R_1) =$$

$$= n\kappa^{n-1} \frac{V}{\kappa^n} = \frac{nV}{\kappa}$$

The expression approaches zero as $\kappa \to \infty$. $\square$

Lemma 3.8 *Let M be a positive integer number* $(M < \infty)$. *Let* $U = \bigcup_{i=1}^{M} I_i$ *be a union of M n-dimensional bounded rectangles, subsets of* $\prod_{i=1}^{n}[\min(T_i), \max(T_i)]$. *Let this the intersection of every two intervals be of zero Lebesgue outer measure. Then*

$$\lim_{\kappa \to \infty} \lambda_n^{\kappa}(U) = \lambda_n^{*}(U)$$

Proof There are maximally $M \cdot 2n\kappa^{n-1}$ distinct tiles R satisfying $R \nsubseteq U$ and $R \cap U \neq \emptyset$.

The difference between $\lambda_m^{\kappa}(U)$ and $\lambda_m^{*}(U) = vol(U)$ is bounded by

$$M \frac{1}{2} \sum vol(R_i) = M \frac{1}{2} 2n \cdot \kappa^{n-1} vol(R_1) =$$

$$= Mn\kappa^{n-1} \frac{V}{\kappa^n} = M \frac{nV}{\kappa}$$

The expression approaches zero as $\kappa \to \infty$. $\square$

Lemma 3.9 *Let X be a (bounded) subset of* $\mathbb{R}^n$ *and let* $J_1, J_2, \ldots$ *be a (countable) sequence of intervals such that* $X \subseteq \bigcup_{j=1}^{\infty} J_j$. *Then there exists a sequence* $I_1, I_2, \ldots$ *of intervals with* $\lambda_n^{*}(I_i \cap I_j) = 0$ *for each pair* $i \neq j$ *such that* $\bigcup_{i=1}^{\infty} I_i = \bigcup_{j=1}^{\infty} J_j$.

Proof The sequence of pairs of intervals from the first sequence is also countable.

Replace every two rectangles that overlap by finitely many new non-overlapping rectangles. □

Lemma 3.10 *Let* $X \subseteq \prod_{i=1}^{n} [\min(T_i), \max(T_i)]$ *be a bounded subset of* $\mathbb{R}^n$ *with lebesgue outer measure* $\lambda_n^*(X) = r \in \mathbb{R}_0^+$.

Let $\varepsilon > 0$. *Let* $J_1, J_2, \ldots$ *be a cover of* X *by intervals such that* $\sum_{i=1}^{\infty} vol(J_i) - r < \varepsilon$.

Let $\delta > 0$ *be a positive real number. Then there exists a number* k_δ *such that* $\sum_{i=1}^{k_\delta} vol(J_i) \in (r - \delta, r + \delta)$.

(That means $\sum_{i=k_\delta}^{\infty} vol(J_i) < \delta$.)

Proof The sequence of partiall sums of the sequence of volumes $(vol(J_i))_{i=1}^{\infty}$ converges to r. That implies the the existence of such k_δ. □

Corollary 3.1 *Let* X *and* δ *be the same as in the previous lemma. Let* $I_1, I_2, \ldots$ *be a cover of* X *by countably many intervals with* $\lambda_n^*(I_i \cap I_j) = 0$ *for each pair* $i \neq j$. *Let* k_δ *be as in the previous lemma. Then there exist two subsets of* X *denoted* X' *and* X'' *such that* $X = X' \cup X''$, $\lambda_n^*(X' \cap X'') = 0$, $X'' \subseteq \bigcup_{i=1}^{k_\delta} I_i$ *and* $X' \subseteq \bigcup_{i=k_\delta}^{\infty} I_i$.

Proof The existence of X' and X'' is obvious (can be defined as $X'' = \bigcup_{i=1}^{k_\delta} I_i \cap X$ and $X' = \bigcup_{i=k_\delta}^{\infty} I_i \cap X$ respectively), there remains the proof of the equality $\lambda_n^*(X' \cap X'') = 0$.

The set $X' \cap X''$ contains only countably many intersections of intervals $I_1, I_2, \ldots$, and these intersection are of measure zero, therefore their union has also measure zero, $\lambda_n^*(X' \cap X'') = 0$. □

Lemma 3.11 *Let* X *be a subset of* $\prod_{i=1}^{n} [\min(T_i), \max(T_i)]$ *with Lebesgue outer measure* $\lambda_n^*(X) = r$. *Then for every* κ *the following statement is true:*

$$\lambda_n^\kappa(X) \leq 2r$$

Proof For any tile J to be counted into the κ-grid measure of a set, the inequalities $\lambda_n^*(X \cap J) \geq \frac{1}{2} vol(J) \Leftrightarrow 2\lambda_n^*(X \cap J) \geq vol(J)$ must hold. Therefore the set X with Lebesgue outer measure r can saturate at most a set of tiles of the overall volume $2r$. □

Theorem 3.3 *Let* $X \subseteq H \in Rect(\mathcal{T})$ *(or more generally* $X \subseteq \prod_{i=1}^{n} [\min(T_i), \max(T_i)]$, *where* $\langle T_1, \ldots, T_n \rangle = \mathcal{T}$ *is the partition of* $\mathcal{B}$*). Then*

$$\lim_{K \to \infty} \lambda_n^\kappa(X) = \lambda_n^*(X). \tag{3}$$

Proof The proof uses the definition of outer Lebesgue measure and our aim is to prove that for every $\varepsilon > 0$ there exists such κ that the difference $|\lambda_n^\kappa(X) - \lambda_n^*(X)| < \varepsilon$.

Let $\varepsilon > 0$ be a positive real number, denote $r = \lambda_n^*(X)$.

For X in the type of interval the proof is easy (Lemma 3.7).

For general X consider sufficiently accurate (with less than $\frac{\varepsilon}{2}$ difference of sum of volumes and r) cover of X with countable collection of intervals whose interiors do not intersect (Lemma 3.9).

There exist two subsets of X denoted X' and X'' such that $X = X' \cup X''$, $\lambda_n^*(X' \cap X'') = 0$, such that the collection of intervals can be divided into a finite part $I_1, \ldots, I_k$ and the remainder $I_{k+1}, \ldots$ such that the sum of volumes of the remainder is sufficiently small (less than $\frac{\varepsilon}{8}$) (Corollary 3.1).

For a bounded set Y with $\lambda_n^*(Y) = s$ the inequality $\lambda_n^\kappa(Y) \leq 2s$ holds for every κ (Lemma 3.11).

For the finite set of k intervals we can find such κ that $|\sum_{i=1}^{k} \lambda_n^\kappa(I_i) - \sum_{i=1}^{k} \lambda_n^*(I_n)| \leq \frac{\varepsilon}{4}$ (applying Lemma 3.7 finitely many times, taking the maximal of obtained values of κ) and the overall difference of $\lambda_n^\kappa(X)$ and $\lambda_n^*(X)$ is less than ε. □

Note that the result applies also to the case with $X \subseteq F \in Facets(H)$ and $\lambda_{n-1}^{\kappa}, \lambda_{n-1}^{*}$.

4   Algorithm

This section introduces procedures for obtaining the reachable state space of the quantitative discrete approximation automaton. Algorithm 1 is a procedure of computing the set of reachable states. Algorithm 2 describes the computation of transitions from one state (i.e. successors) together with their weights using numerical simulations.

The procedure of computing reachable state space (Algorithm 1) is based on breadth first search. States corresponding to initial conditions of the biological system are enqueued first and a list of states already visited is maintained. The computation is always finite, because there are only finitely many possible states of the automaton and each of them can be at most once added and after the computation of its successors removed from the queue.

Algorithm 1 Computing the set of reachable states

Require: $\mathscr{B} = (n, f, \mathscr{T}, \mathscr{I}_C)$, $\kappa \in \mathbb{N}$
Ensure: Reachable = set of all reachable states of the automaton $QDAA_\kappa(\mathscr{B})$
 1: Reachable $\leftarrow \emptyset$
 2: **for all** $H \in \mathscr{I}_C$ **do**
 3:      $s \leftarrow \langle H, H \rangle$
 4:      Reachable $\leftarrow$ Reachable $\cup \{s\}$
 5:      Queue.pushBack(s)
 6: **while** Queue $\neq \emptyset$ **do**
 7:      $s \leftarrow$ Queue.firstElement
 8:      $A \leftarrow$ getSuccessors(s)
 9:      **for all** $a \in A$ **do**
10:          **if** $a \notin$ Reachable **then**
11:              Reachable $\leftarrow$ Reachable $\cup \{a\}$
12:              Queue.pushBack(a)
13: **return** Reachable

Computation of the successors (Algorithm 2) of one state requires determining the rectangles and the entry sets of the successors and weights of the transitions. This can be done approximately using numerical simulations. We sample the entry set of the state and perform numerical simulations with the sampled points as initial conditions and the dynamics of the given biological system as the vector field. For each simulated trajectory we watch whether it leaves the rectangle before given maximal time interval elapses. If this is the case then the location of the exit point through which the trajectory leaves the rectangle is of interest.

Entry sets of the successor states are also determined within Algorithm 2. If the successor is a selfloop state the entry set is empty. For a neighbouring rectangle successor with one common facet the entry set is computed using the exit points locations and more numerical simulations. From the set of exit points in a facet we can estimate the set of κ-tiles of the facet that surely have nonempty intersection with the exit set. It remains to decide in which of the κ-tiles the intersection of the tile with the exit set takes at least one half of the volume of the tile.

To this end we use numerical simulations and the fact that for an autonomous system of ODEs $\dot{x} = f(x)$ with a solution $x(t)$ the function $x(-t)$ is a solution of autonomous system $\dot{x} = -f(x)$. For determining whether to include a κ-tile in the entry set of a successor state, we sample the tile and perform numerical simulations of the trajectories of system $\dot{x} = -f(x)$. If more than one half of the simulated trajectories go through the rectangle and the entry set of the original state, then the κ-tile is included in the entry set of successor state, otherwise the κ-tile is not included.

Weights of the transitions correspond to portions of the set of performed simulations that leave the rectangle to the respective neighbouring rectangles. Weight of the transition from the state to the so-called selfloop state with the same rectangle is determined as the portion of trajectories that do not leave the rectangle in given maximal time interval.

Algorithm 2 Procedure getSuccessors

Require: $\mathscr{B} = (n, f, \mathscr{T}, \mathscr{I})$, $\kappa, M \in \mathbb{N}$, $H \in Rect(\mathscr{T})$, $E \in EntrySets(H)$
Ensure: Successors $= Succs_{\kappa}(\langle H, E \rangle)$
  1: **if** $E = \emptyset$ **then**
  2:      Successors $\leftarrow \{\langle H, \emptyset \rangle\}$
  3:      **return** Successors
  4: $A \leftarrow$ set of M random points in E
  5: ExitPoints $\leftarrow \emptyset$
  6: StaysInside $\leftarrow 0$
  7: **for all** $x_0 \in A$ **do**
  8:      simulate trajectory from x_0 until it leaves H through a point x_1 or given time elapses
  9:      **if** x_1 exists **then**
 10:          ExitPoints $\leftarrow$ ExitPoints $\cup \{x_1\}$
 11:      **else**
 12:          StaysInside $\leftarrow$ StaysInside $+1$
 13: **for all** F $\in Facets(H)$, F$= H \cap H'$ **do**
 14:      **if** ExitPoints $\cap$ F $\neq \emptyset$ **then**
 15:          EntryTiles $\leftarrow \{Z \in Tiles_{\kappa}(F) \mid Z \cap$ ExitPoints $\neq \emptyset\}$
 16:          **for all** $Z \in$ EntryTiles **do**
 17:              $B \leftarrow$ set of M random points in Z
 18:              RealPointsCount $\leftarrow 0$
 19:              **for all** $y_0 \in B$ **do**
 20:                  simulate trajectory from y_0 until it leaves H through a point y_1 or given time elapses
 21:                  **if** $y_1 \in E$ **then**
 22:                      RealPointsCount $\leftarrow$ RealPointsCount $+1$
 23:              **if** RealPointsCount $< \frac{M}{2}$ **then**
 24:                  EntryTiles $\leftarrow$ EntryTiles $\setminus \{Z\}$
 25:      **if** EntryTiles $\neq \emptyset$ **then**
 26:          Successors $\leftarrow$ Successors $\cup \langle H', EntryTiles \rangle$
 27:          Weight$[\langle H, E \rangle][\langle H, EntryTiles \rangle] \leftarrow \dfrac{|\text{ExitPoints} \cap \text{F}|}{|A|}$
 28: **return** Successors

Performing the backward simulations (lines 16–24 of Algorithm 2) can be switched off. The resulting transition system differs from the QDAA in the entry sets, that can be larger. Difference of the outputs can be seen on Figure 5. The algorithm with backward simulations computes the QDAA and for $(\kappa \to \infty)$ approaches the real behaviour of the solutions of dynamics ODE system. On the other hand the algorithm without backward simulations overapproximates the entry sets, therefore the transitions are included even if the entry set of a state is smaller than half of one κ-tile. Both options still lead to automatons with reachable states whose rectangles are included in the set of reachable rectangles of the rectangular abstraction with the same initial rectangles.

The worst case complexity of the algorithms follows. There are at most k^n rectangles in the phase space of the biochemical system, where k is the maximal number of thresholds on one variable. The maximal number of states of QDAA of the form $\langle H, E \rangle$ for a fixed rectangle H is $2n \cdot (2^{\kappa^{n-1}} - 1)$, where n is the dimension of the biochemical system. For the average numbers of visited different states of QDAA with the same rectangle encountered while analysing our evaluation models see the line labeled ρ in Table 1. Complexity of the computation of successors of a given state depends on the dimension of the system, the κ parameter and on the number of simulations M used per one tile. In the worst case when all the tiles are examined (either as a part of entry set or potential exit set) there are $2n \cdot \kappa^{n-1} \cdot M$ simulations.

Visualization of the state space of QDAA involves highlighting the borders of the rectangles H such that there is at least one state $\langle H, E \rangle$ visited during the computation. The intensity of the fill colour of a rectangle H is calculated proportional to the sum of weights of all possible paths from initial set $\mathscr{I}_C$ to the first appearance of states with H as the rectangle. The weight of a finite path is obtained as the product of weights of the subsequent transitions in the path. The sum is always between zero and one.

4.1 Entering and leaving conditions

The computation of successor states proceeds in two steps.

First, the probabilities of potential successors are computed for all κ-tiles of E and summed up to get the probabilities of successors for the whole set E. For each κ-tile of E several numerical simulations of a solution of system (1), with initial point $x(0)$ placed randomly in the tile, are performed. If the computed trajectory satisfies entering and leaving conditions (Definition 4.2) and leaves the box H or the maximal time interval elapses, the number of trajectories leaving H through the particular facet (resp. the number of trajectories assumed to stay forever in H and leading to the transition $\langle H, E \rangle \to \langle H, \emptyset \rangle$) is increased.

Second step of the algorithm takes into account the rectangles $H' \bowtie H$ with nonzero probability of transition $\langle H, E \rangle \to \langle H', \text{yet unknown } E' \rangle$ computed in the first step and determines the entry sets E' of this successors. Denote $F = H' \cap H$. For every κ-tile of F the algorithm decides if the tile is a subset of E'.

Replacing $Entry(F, H)$ with more easily computed set $Entry'(F, H)$ (defined in Definition 4.2) of points satisfying the entering condition in the definition of QDAA does not lead to a nonzero difference in the values of transitions probabilities due to Theorem 4.1.

Definition 4.1 *Let $F \in Facets_i(H)$ be a facet of the form $F = [a_1, b_1] \times \ldots \{c\} \ldots \times [a_n, b_n]$. Define the normal vector $\nu_H(F)$ to the facet F with respect to the rectangle H as the vector $-e_i$ for $c = a_i$ and e_i for $c = b_i$, where e_i denotes the ith vector of the standard basis of $\mathbb{R}^n$ (i.e. the vector orthogonal to F and pointing outside from H).*

Definition 4.2 *Consider rectangle $H, F, F' \in Facets(H)$, a mutiaffine vector field f, a solution $x(t)$ of the system (1) and $r > 0$ satisfying $x(0) \in F, x(r) \in F', \forall t \in (0, r) : x(t) \in H$. We say that $x(t)$ satisfies the entering (resp. leaving) condition with respect to H and f, if $f(x(0)) \cdot \nu_H(F) < 0$ (resp. $f(x(r)) \cdot \nu_H(F') > 0$).*
Define $Entry'(F, H) = \{x \in F | \nu_H(F) \cdot f(x) < 0\}$.

Our next step is to show (in Theorem 4.1) that the focal set of $Entry'(F, H) \setminus Entry(F, H)$ is of measure zero, thus the set difference is insignificant for our volume based notions.

Now we will extend our definition of $Focal(H, E, F')$ set of points from which the trajectories leave the box through a facet F' (Definition 3.6) to $Focal(H, E, X)$ for an arbitrary subset X of facet F'.

Definition 4.3 *Let $E \in EntrySets_\kappa(H), E \neq \emptyset, F' \in Facets(H)$. Let $X \subseteq F'$. Define $Focal(H, E, X)$ the set of all points $y_0 \in E$ such that there exist $\varepsilon, c, \varepsilon' > 0$ and a solution $y(t)$ of the system (1) satisfying $y(t) \in H$ for $t \in (0, c), y(t) \in Inter(H)$ for $t \in (0, \varepsilon), y(c) \in X$, and $y(t) \in Inter(H')$ for $t \in (c, c + \varepsilon')$.*

Theorem 4.1 *Let $E \in EntrySets_\kappa(H), E \neq \emptyset, F' \in Facets(H)$. Let $V_+ = \{x \in ExitSet(H, E, F') | \eta_{F'} \cdot f(x) > 0\}$ and let $V_0 = ExitSet(H, E, F') \setminus V_+$. Then $\lambda_{n'}^*(Focal(H, E, V_0)) = 0$, where n' denotes n for $E = H$, and $(n - 1)$ otherwise.*

Proof First, we have to observe that trajectories from such initial points have to leave the rectangle through a set of measure zero (in fact it is a zero set of a multi-affine polynomial).

Then the proof is simillar to proof of Lemma 3.5. □

Remark 4.1 *Theorem 4.1 implies that replacing $Entry(F,H)$ with $Entry'(F,H)$ in the definition of QDAA (recall that these can be only entry sets of the successor states, not of the initial states $\langle H,H \rangle$), does not lead to a nonzero difference in the values of transitions probabilites.*

Remark 4.2 *Deciding if a point x is an element of $Entry'(F,H)$ is straightforward (compared to checking the ε-condition of Definition 3.2), and $Entry'(F,H) \subseteq Entry(F,H)$.*

Remark 4.3 *Moreover, there is the following symmetry property. Let $F,F' \in Facets(H), x_0 \in F, x_1 \in F'$. There is a solution $y(t)$ of system (1) satisfying $y(0) = x_0, \exists t_1 : y(t_1) = x_1, f(y(0)) \cdot v_H(F) < 0, f(y(t_1)) \cdot v_H(F') > 0$ and $y(t) \in H$ for $t \in (0,t_1)$, if and only if there is a solution $x(t)$ of system $\dot{x}(t) = -f(x)$ satisfying $x(0) = x_1, \exists t_1 : x(t_1) = x_0, -f(x(0)) \cdot v_H(F') < 0, -f(x(t_1)) \cdot v_H(F) > 0$ and $x(t) \in H$ for $t \in (0,t_1)$.*

5   Evaluation and Case Study

In this section the state spaces of several biological models (of dimensions two, four and seven) are explored. Using our prototype implementation of the algorithms from Section 4 implemented in C++, we evaluate our approach on two exemplary biochemical systems. Additionally, we provide a case study held on a biochemical pathway studied in *E. coli* and compare the reachability results of the case study and one of the smaller models with results obtained using the rectangular abstraction approach.

Before we proceed with the models, let us introduce several terms useful for the evaluation. For a biochemical system $\mathscr{B} = \langle n, f, \mathscr{T}, \mathscr{I}_C \rangle$ we denote $\mathscr{R}(\mathscr{I}_C) \subseteq Rect(\mathscr{T})$ the *set of all rectangles reachable from initial set $\mathscr{I}_C$*. For each $H \in Rect(\mathscr{T})$ we denote $mem(H)$ the subset of $\mathscr{R}(\mathscr{I}_C)$ consisting of all states reachable from the initial set with H as rectangle, the so-called *memory of the rectangle H*, $mem(H) = \{\langle R, E \rangle \in \mathscr{R}(\mathscr{I}_C) \mid R = H\}$. Further we denote ρ the average number of memory states (cardinality of $mem(H)$ averaged over all $H \in \mathscr{R}(\mathscr{I}_C)$). The number of QDAA states representing the memory of a rectangle is in the worst case equal to the number of all its possible entry sets. However, the actual values of ρ in our examples are much smaller (see Table 1).

Let us focus on the effect of parameter κ on cardinality of $\mathscr{R}(\mathscr{I}_C)$ and on ρ. Expected behaviour of the approximation is the following. Every facet is divided into κ^{n-1} tiles. A tile is included in the entry set E of some reachable state $\langle H, E \rangle$ if the focal subset $Focal(H,E,A)$ fills at least half of the volume of the tile. For higher values of κ, the set $Tiles_{n'}^{\kappa}(Focal(H,E,A))$ better approximates the set $Focal(H,E,A)$ because of the higher κ-grid resolution. Thus with increasing κ, the quantitative information denoting the probability of reaching states in $\mathscr{R}(\mathscr{I}_C)$ can be computed more precisely. We demonstrate that on models examined below.

5.1   Oscillatory model

First, we consider a 2-dimensional model which is a variant of Lotka-Volterra model with oscillatory behaviour. The oscillatory model has the form of the following multi-affine system:

$$\frac{dX}{dt} = 5 \cdot X - 1 \cdot X \cdot Y$$
$$\frac{dY}{dt} = 0.4 \cdot X \cdot Y - 5.4 \cdot Y$$

We consider the following partition and initial conditions for this model:

$$T_X = \{i | i \in \langle 0, 30 \rangle \subseteq \mathbb{N}_0\}$$
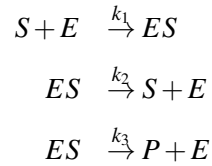$$T_Y = \{i | i \in \langle 0, 12 \rangle \subseteq \mathbb{N}_0\}$$

$$\mathscr{I}_C : X \in \langle 20, 21 \rangle, Y \in \langle 5, 6 \rangle$$

Results achieved on our implementation are presented in Table 1 and visualized in Figure 5. Black rectangles denote the initial set.

5.2 Enzyme kinetics

Similarly, we examined a 4-dimensional model of basic enzyme kinetics based on the following set of reactions:

$$S + E \xrightarrow{k_1} ES$$
$$ES \xrightarrow{k_2} S + E$$
$$ES \xrightarrow{k_3} P + E$$

The corresponding multi-affine ODE model considered in the paper is the following:

$$\frac{dS}{dt} = -0.01 \cdot S \cdot E + 1 \cdot ES$$
$$\frac{dE}{dt} = 1 \cdot ES - 0.01 \cdot S \cdot E + 1 \cdot ES$$
$$\frac{dES}{dt} = -1 \cdot ES + 0.01 \cdot E \cdot S - 1 \cdot ES$$
$$\frac{dP}{dt} = 1 \cdot ES$$

We consider the following partition and initial conditions for this model:

$$T_S = \{0.01, 5, 10, 15, 25, 50, 60, 85, 95, 100\}$$
$$T_E = \{0.01, 5, 10, 15, 25, 50, 60, 85, 95, 100\}$$
$$T_{ES} = \{0.01, 5, 10, 15, 25, 50, 60, 85, 95, 100\}$$
$$T_P = \{0.01, 5, 10, 15, 25, 50, 60, 85, 95, 100\}$$

$$\mathscr{I}_C : S \in \langle 25, 50 \rangle, E \in \langle 95, 100 \rangle, ES \in \langle 0.01, 5 \rangle, P \in \langle 0.01, 10 \rangle$$

Projection of the approximated phase space to the enzyme/substrate plane is shown in Figure 6. For both the oscillatory model and the enzyme kinetics model full version of Algorithm 2 (with backward simulations) was used.

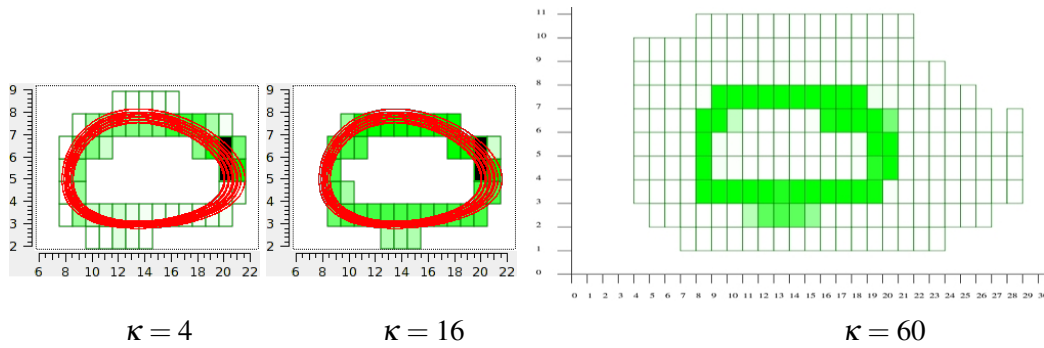$$\kappa = 4 \qquad\qquad \kappa = 16 \qquad\qquad\qquad \kappa = 60$$

Figure 5: Reachability in oscillatory model and comparison with numerical simulation, first two figures were obtained using the full version of Algorithm 2, the third one with lines 16–24 omitted. For comparison: using the rectangular abstraction transition system on this biochemical model, the whole phase space $[0,30] \times [0,12]$ is reachable from the same inital conditions.

	Oscillatory						Enzyme			
κ	4	8	16	32	64	128	4	5	6	7
$|\mathscr{R}(\mathscr{I}_C)|$	52	46	40	39	37	35	76	104	123	166
ρ	1.63	2.2	3.78	2.9	4.57	6	4.36	10.76	16.8	53.6

Table 1: Results for the two models and several different settings of the discretization parameter κ.

$$\kappa = 4 \qquad\qquad \kappa = 6 \qquad\qquad \text{Numerical simulation}$$
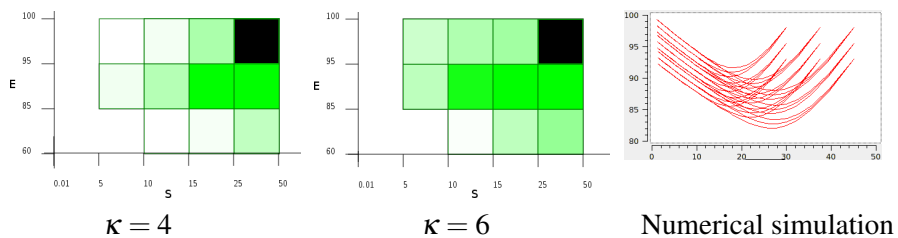
Figure 6: Enzyme kinetics model – projection of the reachable set to the enzyme/substrate plane and comparison with numerical simulation.

5.3   Case Study on E.Coli Ammonium Assimilation Model

We consider a model specifying the ammonium transport from the external environment into cells of *E. Coli* [26]. The model describes the ammonium transport process that takes effect at very low external ammonium concentrations. In such conditions, the transport process complements the deficient ammonium diffusion. The process is driven by a membrane-located ammonium transport protein *AmtB* that binds external ammonium cations NH_4ex and uses their electrical potential to conduct NH_3 into the cytoplasm. In Figure 7, biochemical reactions of this model and the scheme of the transport channel are shown (left and middle).

The level of pH and external ammonium concentration are considered constant.

The system of differential equations:

$$
\begin{aligned}
\frac{d[AmtB]}{dt} &= -k_1[AmtB][NH_4ex] + k_2[AmtB:NH_4] + k_4[AmtB:NH_3] \\
\frac{d[AmtB:NH_3]}{dt} &= k_3[AmtB:NH_4] - k_4[AmtB:NH_3] \\
\frac{d[AmtB:NH_4]}{dt} &= k_1[AmtB][NH_4ex] - k_2[AmtB:NH_4] - k_3[AmtB:NH_4] \\
\frac{d[NH_3in]}{dt} &= k_4[AmtB:NH_3] - k_6[NH_3in][H_{in}] + k_7[NH_4in] + k_9[NH_3ex] \\
\frac{d[NH_4in]}{dt} &= k_6[NH_3in][H_{in}] - k_5[NH_4in] - k_7[NH_4in]
\end{aligned}
$$

Constant species: $NH_3ex, NH_4ex, H_{in}, H_{ex}$.

Initial conditions and threshold numbers:

$$
\begin{aligned}
T_{NH_3ex} &= \{0, 28 \cdot 10^{-9}, 29 \cdot 10^{-9}, 1 \cdot 10^{-5}\} \\
T_{NH_4ex} &= \{0, 49 \cdot 10^{-7}, 5 \cdot 10^{-6}, 1 \cdot 10^{-5}\} \\
T_{AmtB} &= \{0, 1 \cdot 10^{-12}, 1 \cdot 10^{-10}, 5 \cdot 10^{-6}, 9.9 \cdot 10^{-6}, 1 \cdot 10^{-5}\} \\
T_{AmtB:NH_3} &= \{0, 1 \cdot 10^{-7}, 1 \cdot 10^{-5}\} \\
T_{AmtB:NH_4} &= \{0, 1 \cdot 10^{-7}, 1 \cdot 10^{-5}\} \\
T_{NH_3in} &= \{0, 1 \cdot 10^{-8}, 1 \cdot 10^{-7}, 1 \cdot 10^{-6}, 11 \cdot 10^{-7}, 1 \cdot 10^{-5}, 1 \cdot 10^{-4}, 1 \cdot 10^{-3}\} \\
T_{NH_4in} &= \{0, 1 \cdot 10^{-8}, 1 \cdot 10^{-7}, 2 \cdot 10^{-6}, 2.1 \cdot 10^{-6}, 1 \cdot 10^{-6}, 1 \cdot 10^{-5}, 1 \cdot 10^{-4}, 1 \cdot 10^{-3}\}
\end{aligned}
$$

$$
\begin{aligned}
\mathscr{I}_C: \quad & NH_3ex \in \langle 28 \cdot 10^{-9}, 29 \cdot 10^{-9} \rangle, \\
& NH_4ex \in \langle 49 \cdot 10^{-7}, 5 \cdot 10^{-6} \rangle, \\
& AmtB \in \langle 0, 1 \cdot 10^{-5} \rangle, \\
& AmtB:NH_3 \in \langle 0, 1 \cdot 10^{-5} \rangle, \\
& AmtB:NH_4 \in \langle 0, 1 \cdot 10^{-5} \rangle, \\
& NH_3in \in \langle 1 \cdot 10^{-6}, 11 \cdot 10^{-7} \rangle, \\
& NH_4in \in \langle 2 \cdot 10^{-6}, 21 \cdot 10^{-7} \rangle.
\end{aligned}
$$

The upper bounds on concentrations of NH_3in and NH_4in considering the biological system with given initial conditions were estimated as $1.1 \cdot 10^{-6}$ (NH_3in does not exceed the initial concentration) and $5.4 \cdot 10^{-4}$ by the rectangular abstraction (overapproximation).

Reachable intervals using Algorithm 2 without the backward simulations were $[10^{-8}, 1.1 \cdot 10^{-6}]$ for NH_3in (NH_3in does not exceed the initial concentration), and $[10^{-6}, 10^{-5}]$ for NH_4in. This results are in agreement with simulated data and in the case of the concentration of NH_4in the QDAA results are by one order closer to numerical simulations than the rectangular abstraction results as can be seen in the right part of Figure 7.
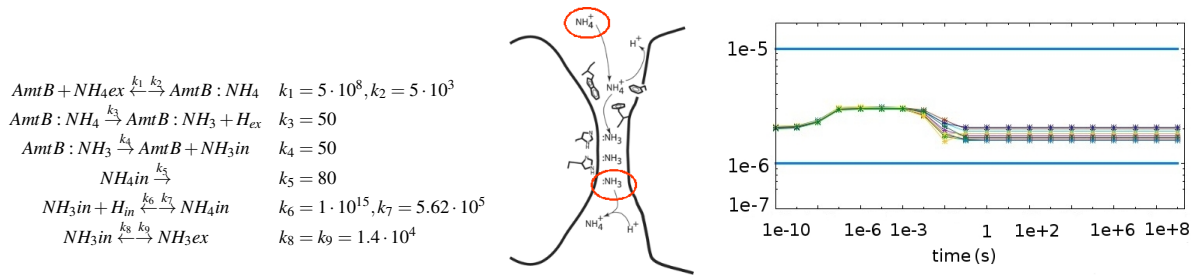
Figure 7: Ammonium transport model (left). Simulations of the ammonium assimilation model from 20 randomly sampled points in $\mathscr{I}_C$ projected on the concentration of NH_4in, blue lines represent bounds on this concentration found by the QDAA - two subsequent thresholds $10^{-6}, 10^{-5}$ (right).

6   Conclusion

We have presented a new theoretical method for finite discrete approximation of autonomous continuous systems equipped with a measure that indirectly quantifies correspondence of the approximated behaviour with the original continuous behaviour. We have provided a computational technique which we implemented in a prototype software. We have examined the implementation on small dimensional models which showed satisfactory results for computing reachability.

The method can be either used as a parameterized simulation technique or employed with rectangular abstraction to quantify the extent of spurious counterexamples. Thus the method can improve the current possibilities of analysis based on model checking techniques. We leave for future work integration of this method into the software for model checking of biochemical dynamical systems [13].

At the theoretical side, we leave for future work precise clarification of our method wrt the rectangular abstraction. From the computational viewpoint, we aim to develop a parallel reachability algorithm that would make the method scalable and applicable to systems of larger dimensions.

References

[1] A. Halász et al. (2007): *Analysis of lactose metabolism in E.coli using reachability analysis of hybrid systems*. *Systems Biology, IET* 1(2), pp. 130 –148.

[2] E. Asarin et al. (2006): *Recent progress in continuous and hybrid reachability analysis*. In: *Computer Aided Control System Design, IEEE International Conference on Control Applications*. pp. 1582 –1587.

[3] G. Batt et al. (2008): *Symbolic reachability analysis of genetic regulatory networks using discrete abstractions*. *Automatica* 44, pp. 982–989.

[4] G. Batt et al. (2008): *Temporal Logic Analysis of Gene Networks under Parameter Uncertainty*. *IEEE Transactions of Automatic Control* 53, pp. 215–229.

[5] H. Kitano et al. (2001): *Foundations of Systems Biology*. The MIT Press.

[6] J. Barnat et al. (2009): *Computational Analysis of Large-Scale Multi-Affine ODE Models*. In: *2009 International Workshop on High Performance Computational Systems Biology (HiBi 2009)*. IEEE Computer Society Press, pp. 81–90.

[7] J. Tyson et al. (1996): *Chemical kinetic theory: understanding cell-cycle regulation*. *Trends in Biochemical Sciences* 21, pp. 89–96.

[8] L. Doyen et al. (2005): *Automatic Rectangular Refinement of Affine Hybrid Systems*. In: *Formal Modeling and Analysis of Timed Systems, LNCS* 3829. Springer Berlin / Heidelberg, pp. 144–161.

[9] P.T. Monteiro et al. (2008): *Temporal Logic Patterns for Querying Qualitative Models of Genetic Regulatory Networks*. In: *ECAI, Frontiers in Artificial Intelligence and Applications* 178. IOS Press, pp. 229–233.

[10] S. Jha et al. (2009): *A Bayesian Approach to Model Checking Biological Systems*. In: *CMSB'09*, LNCS. Springer, pp. 218–234.

[11] G. Antoine & G. Pappas (2006): *Verification Using Simulation*. In: *HSCC'06, LNCS* 3927. Springer, pp. 272–286.

[12] E. Asarin, T. Dang & A. Girard (2007): *Hybridization methods for the analysis of nonlinear systems*. Acta Inf. 43, pp. 451–476.

[13] J. Barnat, L. Brim & D. Šafránek (2010): *High-performance analysis of biological systems dynamics with the DiVinE model checker*. Brief. in Bioinformatics 11, pp. 301–312.

[14] C. Belta & L.C. Habets (2006): *Controlling a class of nonlinear systems on rectangles*. IEEE Transactions on Automatic Control 51(11), pp. 1749–1759.

[15] S. Berman, Á. Halász & V. Kumar (2007): *MARCO: a reachability algorithm for multi-affine systems with applications to biological systems*. In: *HSCC'07*, LNCS. Springer-Verlag, Berlin, Heidelberg, pp. 76–89.

[16] E. M. Clarke, E. A. Emerson & A. P. Sistla (1986): *Automatic verification of finite-state concurrent systems using temporal logic specifications*. ACM Trans. Program. Lang. Syst. 8, pp. 244–263.

[17] T. Dang, C. Le Guernic & O. Maler (2009): *Computing Reachable States for Nonlinear Biological Models*. In: *CMSB'09, LNCS* 5688. Springer Berlin / Heidelberg, pp. 126–141.

[18] M. Feinberg (1987): *Chemical reaction network structure and the stability of complex isothermal reactors I. The deficiency zero and the deficiency one theorems*. Chemical Engineering Science 42, pp. 2229–2268.

[19] J. Fisher & T. A. Henzinger (2007): *Executable cell biology*. Nature biotechnology 25(11), pp. 1239–1249.

[20] L.C. Habets & J. H. van Schuppen (2004): *A control problem for affine dynamical systems on a full-dimensional polytope*. Automatica 40(1), pp. 21 – 35.

[21] Philip Hartman (2002): *Ordinary Differential Equations*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2nd edition.

[22] Morris Hirsch (1976): *Differential Topology*. Springer-Verlag, Berlin.

[23] F. Horn & R. Jackson (1972): *General mass action kinetics*. Archive for Rational Mechanics and Analysis 47, pp. 81–116. 10.1007/BF00251225.

[24] M. Kloetzer & C. Belta (2010): *Reachability analysis of multi-affine systems*. Transactions of the Institute of Measurement and Control 32, pp. 445–467.

[25] F. J. Krambeck (1970): *The mathematical structure of chemical kinetics in homogeneous single-phase systems*. Archive for Rational Mechanics and Analysis 38, pp. 317–347. 10.1007/BF00251527.

[26] H. Ma, F. Boogerd & I. Goryanin (2009): *Modelling nitrogen assimilation of Escherichia coli at low ammonium concentration*. Journal of Biotechnology 144, pp. 175–83.

[27] O. Maler & G. Batt (2008): *Approximating Continuous Systems by Timed Automata*. In: *Proceedings of the 1st international workshop on Formal Methods in Systems Biology*, FMSB '08. Springer, pp. 77–89.

[28] G. F. Oster & A. S. Perelson (1974): *Chemical reaction dynamics*. Archive for Rational Mechanics and Analysis 55, pp. 230–274. 10.1007/BF00281751.

[29] A. Rizk, G. Batt, F. Fages & S. Soliman (2009): *A general computational method for robustness analysis with applications to synthetic gene networks*. Bioinformatics 25, pp. i169–i178.

[30] Walter Rudin (1974): *Real and complex analysis*. McGraw-Hill New York, 2nd edition.