Adaptive Aggregation of Markov Chains: Quantitative Analysis of Chemical Reaction Networks [⋆]

Alessandro Abate[1], Luboš Brim[2], Milan Češka[1,2], and Marta Kwiatkowska[1]

[1] Department of Computer Science, University of Oxford, UK
[2] Faculty of Informatics, Masaryk University, Czech Republic

Abstract. Quantitative analysis of Markov models typically proceeds through numerical methods or simulation-based evaluation. Since the state space of the models can often be large, exact or approximate state aggregation methods (such as lumping or bisimulation reduction) have been proposed to improve the scalability of the numerical schemes. However, none of the existing numerical techniques provides general, explicit bounds on the approximation error, a problem particularly relevant when the level of accuracy affects the soundness of verification results. We propose a novel numerical approach that combines the strengths of aggregation techniques (state-space reduction) with those of simulation-based approaches (automatic updates that adapt to the process dynamics). The key advantage of our scheme is that it provides rigorous precision guarantees under different measures. The new approach, which can be used in conjunction with time uniformisation techniques, is evaluated on two models of chemical reaction networks, a signalling pathway and a prokaryotic gene expression network: it demonstrates marked improvement in accuracy without performance degradation, particularly when compared to known state-space truncation techniques.

1 Introduction

Markov models are widely used in many areas of science and engineering in order to evaluate the probability of certain events of interest. Quantitative analysis of time-bounded properties of Markov models typically proceeds through numerical analysis, via solution of equations yielding the probability of the system residing in a given state at a given time, or via simulation-based exploration of its execution paths. For continuous-time Markov chains (CTMCs), a commonly employed method is uniformisation (also known as Jensen's method), which is based on the discretisation of the original CTMC and on the numerical computation of transient probabilities (that is, probability distributions over time). This can be combined with graph-theoretic techniques for probabilistic model checking against temporal logic properties [4].

There are many situations where highly accurate probability estimates are necessary, for example for reliability analysis in safety-critical systems or for predictive modelling in scientific experiments, but this is difficult to achieve in practice because of the state-space explosion problem. Imprecise values are known to lead to lack of robustness, in the sense that the satisfaction of temporal logic formulae can be affected by small changes to the formula bound or the probability distribution of the model.

Simulation-based analysis does not suffer from this problem and additionally allows dynamic adaptation of the sampling procedure, as e.g. in importance sampling, to the current values of the transient probability distribution. However, this analysis provides only weak precision guarantees in the form of confidence intervals. In order to enable the handling of larger state spaces, two types of techniques have been introduced: state aggregation and state-space truncation. State aggregation techniques build a reduced state space using lumping [6] or bisimulation quotient [21], and have been proposed both in exact [21] and approximate form [10], with the latter deemed more robust than than the exact ones [11]. State-space truncation methods, e.g. fast adaptive uniformisation (FAU) [9,23], on the other hand, only consider the states whose probability mass is not negligible, while clustering states where the probability is less than a given threshold and computing the total probability lost. Unfortunately, though these methods allow the user to specify a desired precision, none provide explicit and general error bounds that can be used to quantify the accuracy of the numerical computation: more precisely, these truncation methods provide a lower bound on the probability distributions in time, and the total probability lost can be used to derive a (rather conservative) upper bound on the (point-wise) approximation error as the sum of the lower bound and of the total probability lost.

Key contributions We propose a novel adaptive aggregation method for Markov chains that allows us to control its approximation error based on explicitly derived error bounds. The method can be combined with numerical techniques such as uniformisation [9,23], typically employed in quantitative verification of Markov chains. The method works over a finite time interval by clustering the state space of a Markov chain sequentially in time, where the quality of the current aggregation is quantified by a number of metrics. These metrics, in conjunction with user-specified precision requirements, drive the process by automatically and adaptively reclustering its state space depending on the explicit error bounds. In contrast to related simulation-based approaches in the literature [13,31] that employ the current probability distribution of the aggregated model to selectively cluster the regions of the state space containing negligible probability mass, our novel use of the derived error bounds allows far greater accuracy and flexibility as it accounts also for the past history of the probability mass within specific clusters.

To the best of our knowledge, despite recent attempts [10,11] the development and use of explicit bounds on the error associated with a clustering procedure is new for the simulation and analysis of Markov chains. The versatility of the method is further enhanced by employing a variety of different metrics to assess the approximation quality. More specifically, we use the following to control the error: (1) the probability distributions in time (namely, the point-wise difference between concrete and abstract distributions), (2) the time-wise likelihood of events (L_1 norm and total variation distance), as well as (3) the probability of satisfying a temporal logic specification.

We implement our method in conjunction with uniformisation for the computation of probability distributions of the process in time, as well as time-bounded probabilities (a key procedure for probabilistic model checking against temporal logic specifications), and evaluate it on two case studies of chemical reaction networks. Compared to fast adaptive uniformisation as implemented in PRISM [9], currently the best performing technique in this setting, we demonstrate that our method yields a marked improvement in numerical precision without degrading its performance.

Related work (Bio-)chemical reaction networks can be naturally analysed using discrete stochastic models. Since the discrete state space of these models can be large or even infinite, a number of numerical approaches have been proposed to alleviate the associated state-space explosion problem. For biochemical models with large populations of chemical components, fluid (mean-field) approximation techniques can be applied [5] and extended to approximate higher-order moments [12]: these deterministic approximations lead to a set of ordinary differential equations. In [16], a hybrid method is proposed that captures the dynamics using a discrete stochastic description in the case of small populations and a moment-based deterministic description for large populations. An alternative approach assumes that the transient probabilities can be compactly approximated based on quasi product forms [3]. All the mentioned methods do not provide explicit accuracy bounds of approximation.

A widely studied model reduction method for Markov models is state aggregation based on lumping [6] or (bi-)simulation equivalence [4], with the latter notion in its exact [21] or approximate [10] form. In particular, approximate notions of equivalence have led to new abstraction/refinement techniques for the numerical verification of Markov models over finite [11] as well as uncountably-infinite state spaces [1,2,26]. Related to these techniques, [27] presents an algorithm to approximate probability distributions of a Markov process forward in time, which serves as an inspiration for our adaptive scheme. From the perspective of simulations, adaptive aggregations are discussed in [13] but no precision error is given: our work differs by developing an adaptive aggregation scheme, where a formal error analysis steers the adaptation.

An alternative method to deal with large/infinite state spaces is truncation, where a lower bound on the transient probability distribution of the concrete model is computed, and the total probability mass that is lost due to this truncation is quantified. Such methods include finite state projections [24], sliding window abstractions [18], or fast adaptive uniformisation (FAU) [9, 23]. Apart from truncating the state space by neglecting states with insignificant probability mass, FAU dynamically adapts the uniformisation rate, thus significantly reducing the number of uniformisation steps [30]. The efficiency of the truncation techniques depends on the distribution of the significant part of the probability mass over the states, and may result in poor accuracy if this mass is spread out over a large number of states, or whenever the selected window of states does not align with a property of interest.

Summarising, whilst a number of methods have been devised to study or to simulate complex biochemical models, in most cases a rigorous error analysis is missing [13, 22, 31], or the error analysis cannot be effectively used to obtain accurate bounds on the probability distribution or on the likelihood of events of interest [17].

Structure of this article Section 2 introduces the sequential aggregation approach to approximate the transient probability distribution (that is, the distribution over time) of discrete-time Markov chains, and quantifies bounds on the introduced error according to three different metrics. Section 3 applies the aggregation method for temporal logic verification of Markov chains. In Section 4, we implement adaptive aggregation for continuous-time Markov chain models of chemical reaction networks, in conjunction with known techniques such as uniformisation and threshold truncation. Finally, Section 5 discusses experimental results.

2 Computation of the transient probability distribution

We first work with discrete-time labelled Markov chains (LMC), and in Section 4 we show how to apply the obtained results to (labelled) continuous-time Markov chains. Formally, an LMC is defined as a triple (S, P, L), where

- $S = \{s_1, \ldots, s_n\}$ is the finite state space of size n;
- $P : S \times S \rightarrow [0, 1]$ is the transition probability matrix, which is such that $\forall j \in S$: $\sum_{i=1}^n P_{ji} = \sum_{i=1}^n P(j, i) = 1$;
- $L : S \rightarrow 2^\Sigma$ is a labelling function, where Σ is a finite alphabet built from a set of atomic propositions.

Whenever clear from the context, we refer to the model simply as (S, P). The model is initialised via distribution $\pi_0 : S \rightarrow [0, 1], \sum_{s \in S} \pi_0(s) = 1$, and its transient probability distribution at time step $k \geq 0$ is

$$\pi_{k+1}(s) = \sum_{s' \in S} \pi_k(s')P(s', s), \tag{1}$$

or more concisely as $\pi_{k+1} = \pi_k P$ (where the π_k's are row vectors). We are interested in providing a compact representation and an efficient computation of the vectors π_k.

Sequential aggregations of the Markov chain

Consider the finite time interval of interest $[0, 1, \ldots, N]$. Divide this interval into a given number (q) of sub-intervals, namely select $N_1, N_2, \ldots, N_q : \sum_{i=1}^q N_i = N$, and consider the evolution of the model within the corresponding l-th interval $[\sum_{i=0}^{l-1} N_i, \sum_{i=0}^l N_i]$, for $l = 1, \ldots q$, and where we have set $N_0 = 0$.

We assume that a specific state-space aggregation is given, for each of the q sub-intervals of time. Later, in Section 4, we show how such aggregations can be obtained adaptively, based on a number of measures (such as the current value of the aggregated transient probability distribution, or the accrued aggregation error in time). In particular, at the l-th step (where $l = 1, \ldots, q$), the state space is partitioned (clustered) as $S = \cup_{i=1}^{m_l} S_i^l$ (consider that the cardinality index m_l has been reasonably selected so that $m_l << n$), and denote the *abstract (aggregated) state space* simply as S^l and its elements (the *abstract states*) with $\phi_i, i = 1, \ldots, m_l$. Introduce abstraction and refinement maps as $\alpha^l : S \rightarrow S^l$ and $A^l : S^l \rightarrow 2^S$, respectively – the first takes concrete points into abstract ones, whereas the latter relates abstract states to concrete partitioning sets. For any pair of indices $i, j = 1, \ldots, m_l$, define the abstract transition matrix as

$$P^l(\phi_j, \phi_i) \doteq \frac{1}{|A^l(\phi_i)|} \sum_{s \in A^l(\phi_i)} \sum_{s' \in A^l(\phi_j)} P(s', s).$$

The intuition behind the aggregated matrix P^l is that it encompasses the average *incoming* probability from clusters S_j to S_i. The shape of this matrix is justified by the structure of the update equation in (1). Given the aggregated Markov chain, we shall work, for all $s \in S^l$, with the following recursions:

$$\pi_{k+1}^l(s) = \sum_{s' \in S^l} \pi_k^l(s')P^l(s', s).$$

The smaller, aggregated model (S^l, P^l) serves as basis for an approximate computation of the transient probability in time: we now calculate an explicit upper bound on the approximation error. In order to quantify this error, we define a function $\epsilon^l : [1, \ldots, m_l]^2 \to [0, 1]$, as follows:

$$\epsilon^l(j, i) \doteq \max_{s \in S_i^l} \left| \frac{|S_i^l|}{|S_j^l|} P(S_j, s) - P^l(\phi_j, \phi_i) \right|. \tag{2}$$

Intuitively, this quantity accounts for the difference between the average incoming probability between a pair (j, i) of partitioning sets, and the worst-case (rescaled) point-wise incoming probability between those two sets. Introduce the terms $\epsilon^l(j) := \sum_{i=1}^{m_l} \epsilon^l(j, i)$.

Finally, define, for all $s \in S$, $\tilde{\pi}_k^l(s) = \pi_k^l(\alpha^l(s))/|A^l(\alpha^l(s))|$ as a (normalised) piecewise constant approximation of the abstract functions π_k^l. Functions $\tilde{\pi}_k^l$, being defined over the concrete state space S, will be employed for comparison with the original distribution functions π_k. Specifically, for the initial interval $[N_0, N_1]$ (with $l = 1$), approximate the initial distribution π_0 by π_0^1 as: $\forall s \in S^1, \pi_0^1(s) = \sum_{s' \in A^1(s)} \pi_0(s')$. Similarly, we have that $\forall s \in S, \tilde{\pi}_0^1(s) = \pi_0^1(\alpha^1(s))/|A^1(\alpha^1(s))|$.

Remark 1. Exact and approximate probabilistic bisimulations [10, 21] build a quotient or a cover of the state space of the original model based on matching or approximating the "outgoing probability" from concrete states – for example, exact probabilistic bisimulation compares, for state pairs (s_1, s_2) within a partition, the "outgoing" probabilities $P(s_1, B)$ and $P(s_2, B)$ over partitions B. On the other hand, in (2) we approximate the "incoming probability", as motivated by the approximation of the recursions in (1). $\square$

Explicit error bounds for the quality of the sequential aggregations

Let us consider the aggregated model (S^1, P^1) (for $l = 1$) and, given the aggregated vector π_0^1, the time-wise updates $\pi_{k+1}^1 = \pi_k^1 P^1, k = N_0, \ldots, N_1 - 1$. Introduce the interpolated vectors $\tilde{\pi}_{k+1}^1(s), s \in S$, defined as $\tilde{\pi}_{k+1}^1(s) = \pi_{k+1}^1(\alpha^1(s))/|A^1(\alpha^1(s))|$. We are interested in a bound on the point-wise error defined over the concrete state space, namely $\forall s \in S, k = N_0, \ldots, N_1, \left|\pi_k(s) - \tilde{\pi}_k^1(s)\right|$, or equivalently a bound for $\left|\pi_k(s) - \frac{\pi_k^1(\alpha^1(s))}{|A^1(\alpha^1(s))|}\right|$. Such a point-wise bound directly allows for expressing a global error for the infinity norm of the difference between the two distribution vectors, namely

$$\left\|\pi_k - \tilde{\pi}_k^1\right\|_\infty = \max_{s \in S} \left|\pi_k(s) - \tilde{\pi}_k^1(s)\right|.$$

Beyond the first aggregation ($l = 1$), the next statement explicitly characterises such a bound over the entire sequence of q re-aggregations and the time interval $[0, 1, \ldots, N]$.

Proposition 1. *Consider a sequential q-step aggregation strategy, characterised by times $N_l : \sum_{l=1}^q N_l = N$, partitions $S = \cup_{i=1}^{m_l} S_i^l$, and matrices P^l. We obtain*

$$\left|\pi_N(s) - \tilde{\pi}_N^q(s)\right| \le c(s)^N \left|\pi_0(s) - \tilde{\pi}_0^1(s)\right|$$

$$+ \sum_{l=1}^q c(s)^{N - \sum_{i=0}^l N_i} \left\{ \frac{1}{|A^l(\alpha^l(s))|} \sum_{j=1}^{m_l} \epsilon^l(j, \alpha^l(s)) \sum_{k=0}^{N_l - 1} \pi_{\sum_{i=0}^{l-1} N_i + k}^l(j) + \gamma_{l-1}^l(s) \right\},$$

where we have set $c(s) = P(S, s)$, *and where* $\gamma_{l-1}^l(s) = \left| \tilde{\pi}_{\sum_{i=0}^{l-1} N_i}^{l-1}(s) - \tilde{\pi}_{\sum_{i=0}^{l-1} N_i}^l(s) \right|$ *for*
$l = 1, \ldots, q$, *with* $\gamma_0^1(s) = 0$, $\forall s \in S$.

Remark 2. A few comments on the structure of the error bounds are in order. The overall error is composed of two main contributions, one depending on the error accrued within single aggregation steps, and the other ($\gamma_{l-1}^l(s)$) depending on the q re-aggregations (that is, an update from the current partition to the next).

The first term of the first contribution further depends on the point-wise error in the distributions initialised at each aggregation, namely, $\left| \pi_{\sum_{i=0}^l N_i}(s) - \tilde{\pi}_{\sum_{i=0}^l N_i}^l(s) \right|$: this quantity, discounted by the N_l-th power of the factor $c(s)$ (accounting for contractive or expansive dynamics), builds up recursively to yield the global (over the q aggregation steps) quantity $c(s)^N \left| \pi_0(s) - \tilde{\pi}_0^1(s) \right|$. The second term of the first contribution, on the other hand, accounts for the error due to the approximation of the transition probability matrix (terms ϵ^l), averaged over the accrued running distribution functions (factors π^l).

The intuition on factor $c(s)$ is the following: if the model is "contractive" (in a certain probabilistic sense) towards a point s, the factor $c(s)$ is likely to be greater than one; on the other hand, if the distribution in time is "dispersed," then it is likely that $c(s) < 1$ over a large subset of the state space. The quantity $c(s) = P(S, s)$ might be decreased if we work on a subset of S: this might happen with a discrete-time chain obtained from a corresponding continuous-time model via FAU [9, 23], or through the interaction of the factor $c(s)$, $s \in S$, with atomic propositions defined specifically over subsets of the state space S. □

Corollary 1. *Consider the same setup as in Proposition 1. A bound for the quantity* $\left\| \pi_N - \tilde{\pi}_N^q \right\|_\infty$ *can be obtained from that in Proposition 1 by straightforward adaptation and setting* $c = \max_{s \in S} c(s)$, *and* $\gamma_{l-1}^l = \max_{s \in S} \gamma_{l-1}^l(s)$, $l = 1, \ldots, q$.

In addition to point-wise errors, we seek a bound for the following global error,

$$\left\| \pi_k - \tilde{\pi}_k^1 \right\|_1 = \sum_{s \in S} \left| \pi_k(s) - \tilde{\pi}_k^1(s) \right|, \quad \forall k = 0, \ldots, N_1,$$

and its further extension to successive aggregations and time steps $k = N_1 + 1, \ldots, N$. This L_1-norm measure is related to the "total variation distance" over events in the σ-algebra 2^S at each time step k. This measure is commonly used in related literature [8, 29], and refers to differences in probability of events defined over sets in S at a specific point in time k. The corresponding error bound is explicitly quantified as follows.

Proposition 2. *Consider a q-step sequential aggregation strategy characterised by the times* $N_l : \sum_{l=1}^q N_l = N$, *partitions* $S = \cup_{i=1}^{m_l} S_i^l$, *and matrices* P^l. *We obtain*

$$\left\| \pi_N - \tilde{\pi}_N^q \right\|_1 \le \left\| \pi_0 - \tilde{\pi}_0^1 \right\|_1 + \sum_{l=1}^q \left\{ \sum_{j=1}^{m_l} \epsilon^l(j) \sum_{k=0}^{N_l-1} \pi_{\sum_{i=0}^{l-1} N_i + k}^l(j) + \Gamma_{l-1}^l \right\},$$

where for $l = 1, \ldots, q$, $\Gamma_{l-1}^l = \left\| \tilde{\pi}_{\sum_{i=0}^{l-1} N_i}^{l-1} - \tilde{\pi}_{\sum_{i=0}^{l-1} N_i}^l \right\|_1$, *and where we have set* $\Gamma_0^1 = 0$.

3   Aggregations for model checking of time-bounded specifications

In Section 2, we have introduced a sequential aggregation procedure to approximate the computation of the transient probability distribution of a Markov chain. The derived bounds allow for a comparison of aggregated and concrete models either point-wise, or according to a global measure of the differences in the probability of events over the state space, at a specific point in time. We now show how to employ the aggregation method for quantitative verification against probabilistic temporal logics such as PCTL. We focus on a bounded variant of the probabilistic safety (invariance) property, which corresponds to time-bounded invariance for continuous-time Markov chains.

Consider the LMP (S, P, L). We focus on properties expressed over the atomic propositions AP, namely the set of finite strings over the labels 2^{AP}, and on how to approximately compute the likelihood associated to such strings. In particular, consider a step-bounded safety formula [4], namely $\mathbb{P}_{=?}\left(G^{\leq N}\Phi\right)$, where $N \in \mathbb{N}$, and $\Phi \in 2^{\Sigma}$, $\text{Sat}(\Phi) \subseteq S$[3]. This specification expresses the likelihood that a trajectory, initialised according to a distribution (say, π_0) over the state space S, resides within set Φ over the time interval $[0, 1, \ldots, N]$. The specification of interest can be characterised as follows: for any $s \in S, k = 0, 1, \ldots, N - 1$,

$$V_0(s) = 1_{\Phi}(s)\pi_0(s), \qquad V_{k+1}(s) = 1_{\Phi}(s) \sum_{s' \in S} V_k(s')P(s', s),$$

so that $\mathbb{P}_{=?}\left(G^{\leq N}\Phi\right) = \sum_{s \in S} V_N(s)$. It is well known that the computed quantity depends on the choice of the initial distribution π_0 (which can in particular be a point mass for a distinguished initial state). As should be clear from the recursion above (use of indicator functions 1_{Φ}), it is sufficient to restrict the recursive updates to within the set of points labelled by Φ.

As before, consider the global finite interval $[0, 1, \ldots, N]$, and divide it via intervals of duration $N_1, N_2, \ldots, N_q : \sum_{i=1}^{q} N_i = N$. Specifically, for the initial interval $[N_0, N_1]$ (corresponding to index $q = 1$), partition set Φ as $\Phi = \cup_{i=1}^{m_1} \Phi_i^1$ – notice that the partition does not cross the boundaries of the set Φ. Thus $S^1 = \Phi^1 \cup \{a_1\} = \{1, \ldots, m_1, a_1\}$, where a_1 is associated with the complement set $S \setminus \Phi$. Introduce abstraction and refinement maps as $\alpha^1 : S \to S^1$ and $A^1 : S^1 \to 2^S$, the abstract transition matrix P^1, and function $\epsilon^1 : [1, \ldots, m_1]^2 \to [0, 1]$ as

$$\epsilon^1(j, i) \doteq \max_{s \in \Phi_i^1} \left| \frac{|\Phi_i^1|}{|\Phi_j^1|} P(\Phi_j^1, s) - P^1(\phi_j, \phi_i) \right|.$$

Further, approximate π_0 as: $\forall s \in S^1, \pi_0^1(s) = \sum_{s' \in A^1(s)} \pi_0(s')$. Introduce, $\forall s \in S^1$, cost functions V_i via the following recursions:

$$V_0^1(s) = 1_{\Phi^1}(s)\pi_0^1(s), \qquad V_{k+1}^1(s) = 1_{\Phi^1}(s) \sum_{s' \in S^1} V_k^1(s')P^1(s', s),$$

and, $\forall s \in S, \tilde{V}_k^1(s) = V_k^1(\alpha^1(s))/|A^1(\alpha^1(s))|$, as a (normalised) piecewise constant approximation of the abstract functions V_k^1, and in particular initialised as $\tilde{\pi}_0^1(s) =$

---

[3] For the sake of simplicity, we shall often loosely identify the set $\text{Sat}(\Phi)$ with label Φ.

$\pi_0^1(\alpha^1(s))/|A^1(\alpha^1(s))|$. We shall derive explicit bounds on the computation of the error:

$$\left| \sum_{s \in \Phi} V_{N_1}(s) - \sum_{s \in \Phi} \tilde{V}_{N_1}^1(s) \right| = \left| \sum_{i=1}^{m_1} \sum_{s \in \Phi_i} \left(V_{N_1}(s) - \tilde{V}_{N_1}^1(s) \right) \right|,$$

and extend them over successive aggregation and time steps $k = N_1 + 1, \ldots, N$. Notice that, in this instance, we are comparing two scalars, comprising the likelihoods associated with the specification of interest computed over the concrete and abstract models, respectively. More precisely, in general we have:

$$\left| \sum_{s \in \Phi_i} V_{N_1}(s) - \sum_{s \in \Phi_i} \tilde{V}_{N_1}^1(s) \right| = \left| \sum_{s \in \Phi_i} V_{N_1}(s) - \sum_{s \in \Phi_i} \frac{V_{N_1}^1(\alpha^1(s))}{|A^1(\alpha^1(s))|} \right| = \left| \sum_{s \in \Phi_i} V_{N_1}(s) - V_{N_1}^1(i) \right|.$$

Proposition 3. *Consider a q-step sequential aggregation strategy characterised by corresponding times* $N_l : \sum_{l=1}^q N_l = N$, *partitions* $\Phi = \cup_{i=1}^{m_l} \Phi_i^l$, *and matrices* P^l. *We obtain:*

$$\left| \sum_{s \in \Phi} V_N(s) - \sum_{s \in \Phi} \tilde{V}_N^q(s) \right| \le \sum_{l=1}^q \sum_{i=1}^{m_l} \epsilon^l(i) \sum_{k=0}^{N_l-1} V_{\sum_{i=0}^{l-1} N_i + k}^l(i).$$

Remark 3. We give some intuition regarding the structure of the bounds. The quantity depends on a summation over q aggregation steps. It expresses the accrual of the error incurred over the outgoing probability from the i-th partition (term $\epsilon^l(i)$), averaged over the history of the cost function over that partition. Note the symmetry between the shape of the bound and the recursive definition of the quantities of interest. □

4 Quantitative analysis of chemical reaction networks

A chemical reaction network describes a biochemical system containing M chemical species participating in a number of chemical reactions. The state of a model of the system at time $t \in \mathbb{R}^+$ is the vector $\mathbf{X}(t) = (X_1(t), X_2(t), \ldots, X_M(t))$, where X_i denotes the number of molecules of the i-th species [15]. Whenever a single reaction occurs the state changes based on the stoichiometry of the corresponding reaction. We use S to denote the set of (discrete) states. Further, for $s \in S$, $\pi_t(s)$ denotes the probability $\mathbb{P}(\mathbf{X}(t) = s)$. Assuming finite volume and temperature, the model can be interpreted as a continuous-time Markov chain (CTMC) $C = (S, R)$, where the rate matrix $R(s, s')$ gives the rate of a transition from states s to s', and π_0 specifies the initial distribution over S. The time evolution of the model is governed by the Chemical Master Equation (CME) [15], namely $\frac{d}{dt}\pi_t = \pi_t \cdot Q$, where Q is the infinitesimal generator matrix, defined as $Q(s, s') = R(s, s')$ if $s \ne s'$, and as $1 - \sum_{s'' \ne s} R(s, s'')$ otherwise. The exact solution of the CME is in general intractable, which has led to a number of possible numerical approximations [25]. We employ uniformisation [30], which in many cases outperforms other methods and also provides an arbitrary, user-defined approximation precision.

Uniformisation is based on a time-discretisation of the CTMC. The distribution π_t is obtained as a sum (over index i) of terms giving the probability that i discrete reaction steps occur up to time t: this is a Poisson random variable $\gamma_{i,\lambda \cdot t} = e^{-\lambda \cdot t} \cdot \frac{(\lambda \cdot t)^i}{i!}$,

where the time delay is exponentially distributed with rate λ. More formally, $\pi_t = \sum_{i=0}^{\infty} \gamma_{i,\lambda \cdot t} \cdot \pi_0 \cdot \tilde{Q}^i \approx \sum_{i=0}^{N} \gamma_{i,\lambda \cdot t} \cdot \pi_0 \cdot \tilde{Q}^i$, where $\tilde{Q}$ is the uniformised infinitesimal generator matrix defined using terms $\frac{R(s,s')}{\lambda}$, and where the uniformisation constant λ is equal to the maximal exit rate $\sum_{s'' \neq s} R(s, s'')$. Although the sum is in general infinite, for a given precision an upper bound N can be estimated using techniques in [14], which also allow for efficient computation of the Poisson probabilities $\gamma_{i,\lambda \cdot t}$.

For complex models with very large or possibly infinite state spaces, the above numerical approximations are computationally infeasible, and are typically combined with (dynamical) state-space truncation methods, such as finite state projection [24], sliding window abstraction [18], or fast adaptive uniformisation [9, 23] (FAU). The key idea of these truncation methods is to restrict the analysis of the model to a subset of states containing significant probability mass. One can easily compute the probability lost at each uniformisation step and thus obtain the total probability lost by truncation. As such, these truncation methods provide a lower bound on the quantities π_t, and the quantified probability lost can be used to derive a (rather conservative) upper bound on the approximation error: the sum of the lower bound and the probability lost gives an upper bound for the point-wise error. Moreover, a (pessimistic) bound on the L_1-norm over a general subset of the state space is obtained by multiplying the probability lost by the number of states in the concrete subset.

Adaptive aggregation for CTMC models of chemical reaction networks

The aggregation methods in the previous sections can be directly applied to uniformised CTMCs, such as those arising from chemical reaction networks. We now discuss how the aggregation unfolds sequentially in time and how the derived error bounds can be used for the aggregation method in this setting.

Recall from eq. (2) that the derivation of the error bounds for the aggregation procedure requires a finite state space: for *infinite-state* CTMCs, the aggregation method can be combined with state-space truncation (alongside time uniformisation), in order to accelerate computations in cases where the set of significant states is still too large. On the other hand, for *finite-state* CTMC models, adaptive aggregations can be regarded as an orthogonal strategy to truncation, and can be directly applied in conjunction with time uniformisation. In order to compare the precision and reduction capability of our method to that of FAU, we thus assume that the population of each species is bounded, which ensures fairness of experimental evaluation.

The key ingredient of the proposed aggregation method is a partitioning strategy that controls and adapts the clustering of the state space over the given finite time interval. Algorithm 1 summarises the scheme for transient probability calculation (the adaptive aggregation for an invariant property as in Section 3 unfolds similarly). The procedure starts with a given partition S^1 of the state space S (obtained by the procedure initAggregation on line 2). It dynamically (and automatically) updates the current partition when needed, thus providing new abstract state spaces S^l over the l-th time interval $[\sum_{i=0}^{l-1} N_i, \sum_{i=0}^{l} N_i]$, where $l = 2, \ldots, q$ and $q \ll N$. The update of the current l-th clustering is performed after N_l steps, that is, whenever the error accrual exceeds a threshold ensuring the user-defined precision θ (function checkAggregation on line 6). At the same time, the aggregation strategy aims to minimise the average size of the

---

Algorithm 1 Adaptive aggregation scheme for computation of transient probability

---

Require: Finite CTMC $C = (S, R)$, initial distribution π_0, time t, and bound θ on L_1-norm error
Ensure: $globalError \leq \theta$

 1: $(P, N) \leftarrow$ uniformise (C, t) ; $l \leftarrow 1$
 2: $\left(S^l, P^l, \pi_0^l, \epsilon^l\right) \leftarrow$ initAggregation (S, P, π_0)
 3: **for** $(k \leftarrow 0; k \leq N, k \leftarrow k + 1)$ **do**   ▷ perform N uniformisation steps
 4:     $(globalError, AccumErrors) \leftarrow$ computeErrors $\left(\pi_k^l, \epsilon^l, k\right)$
 5:     $\pi_{k+1}^l(s) = \sum_{s' \in S^1} \pi_k^1(s') P^l(s', s)$ ▷ update the probability distribution
 6:     **if** checkAggregation $\left(\epsilon^l, \pi_{k+1}^l, AccumErrors, \theta\right)$ = false **then**
 7:         $(S^{l+1}, P^{l+1}, \pi_{k+1}^{l+1}, \epsilon^{l+1}) \leftarrow$ Recluster $\left(S^l, P^l, \pi_{k+1}^l, \epsilon^l, AccumErrors\right)$
 8:         $AccumErrors \leftarrow 0; l \leftarrow l + 1$

---

abstract state space, defined as $avg = \sum_{l=0}^{q} N_l \cdot |S^l|/N$. We consider two adaptive strategies, one time-*local* and the other *history*-dependent, both of which are driven by the shape of the derived error bounds – in particular, the history-dependent strategy exactly employs the calculated error bounds. Both strategies are parametrised by thresholds, which ensure the required overall precision θ and account for the size of the concrete state space as well as the number of uniformisation steps N.

The *history*-dependent strategy is based on the available history contributing to the shape of the derived errors: for the l-th aggregation step and the given i-th cluster of the current partition, it tracks the sum of the errors accumulated in the interval $[\sum_{i=0}^{l-1} N_i, \sum_{i=0}^{l-1} N_i + k]$ for $k = 1, \ldots, N_l$, according to the explicit bounds derived in Section 2 (line 4 of Algorithm 1). At each step k, the obtained value (averaged over k steps) reflects the (averaged) error accrual for each cluster (array *AccumErrors*) and is used to drive the partitioning procedure.

The function checkAggregation determines (using *AccumErrors*) if the current clustering meets the desired threshold, or if a refinement is desirable: during re-clustering, a locally coarser abstraction may as well be suggested by merging clusters. The function Recluster provides the new clustering based on the error bounds, which are functions of *AccumErrors*, of the local contributions ϵ^l, and of the (history of the) distribution π_k^l (or of the cost V_k^l in the case of safety verification). In contrast to the adaptive method presented in [13] and based exclusively on local heuristics, our strategy closely reflects the shape of the derived, history-dependent error bounds. Note that the aggregation strategy applied to chemical reaction networks aligns well with the known structure of the underlying CTMCs. In particular, the state-space clustering employs the spatial locality of the distribution of transitions in the M-dimensional space [13, 31] (M is the number of chemical species), usually leading to relatively uniform probability mass over adjacent states and thus to strategies that cluster neighbouring states.

A simpler re-clustering strategy (denoted in the experiments as *local*) employs at each uniformisation step k only the product of the local error ϵ^l with the probability distribution π_k^l (or with the cost function V_k^l). In other words, a local re-clustering is performed if the local error depending on $\epsilon^l \pi_k^l$ (respectively, on $\epsilon^l V_k^l$) is above a given threshold. This intuitive scheme is similar to the local heuristic employed in [13].

We will show that the history-based strategy is more flexible with respect to the required precision and aggregation size. Our experiments confirm that, while based on

error bounds that over-estimate the actual empirical error incurred in the aggregation, the history-based strategy tends to outperform the more intuitive and easier local strategy, with respect to key performance metrics affecting the practical use of the adaptive aggregation. This shows that the computed errors not only serve as a means to certify the accuracy of approximation, but can also be used to effectively drive the aggregation procedure. In particular, the metrics we are interested in are: (1) the value of *avg* representing the state-space reduction; (2) the accuracy of the empirical results of the abstract model; (3) the total number of re-clusterings; and (4) the actual value of the error bounds (compared to the empirical errors).

The number of re-clusterings (denoted by q) is crucial for the performance of the overall scheme, since each re-clustering requires $O(|S| + |P|)$ steps, which is similar to performing a few uniformisation steps for the concrete model. As such, the number of re-clusterings should be significantly smaller than the total number of uniformisation steps. Therefore, in our experiments we use thresholds that favour fewer re-clusterings over coarser abstractions. Finally, note that the adaptive aggregation scheme can be combined with the adaptive uniformisation step as well as with dynamic state-space truncation [9, 23, 30], which updates the uniformisation constant λ for different time intervals, thus decreasing the number of overall uniformisation steps N.

Illustrative example We resort to a two-dimensional discrete Lotka-Volterra "predator-prey" model [15] to illustrate the history-dependent aggregation strategy. The maximal population of each species is bounded by 2000, thus the concrete model has 4M states. The initial population is set to 200 predators and 400 preys.

Figure 1 displays the outcome of the adaptive procedure (top row) at three distinct time steps and (bottom row) the current probability distribution of the concrete model. For ease of visualisation, the top plots display for each point of the concrete model the size of its corresponding cluster, where we have limited the maximal size to 100 states. Note the close correspondence between the error bounds and the computed empirical errors, and the limited number of re-clusterings needed (one in about 200 uniformisation steps). Observe that the single-state clusters (red colour in the plots) tend to collect where the current probability distribution peaks. The figure also illustrates a memory effect due to the history-dependent error bounds employed by the aggregation.

5   Experimental evaluation on two case studies

We have developed a prototype implementation of the adaptive aggregation for the quantitative analysis of chemical reaction networks modelled in PRISM [20]. We have evaluated the scheme on two case studies in comparison with FAU [9] as implemented in the explicit engine of PRISM. In order to ensure comparability between the two schemes, which employ different data structures, rather than measuring execution time we have focused on assessing performance based on measures that are independent of implementation, and specifically focused on the metrics (1)-(4) introduced in Section 4 (model reduction, empirical accuracy, number of re-clusterings, and formal error bounds). For the same reason, we have not incorporated heuristics such as varying the maximal cluster size, optimally selecting error thresholds, or use of advanced clustering methods, which can be employed to further optimise the adaptive scheme.

Number of states:	112K	218K	350K
Epirical L_1 norm:	1.41E-7	3.27E-7	6.82E-7
Formal bounds:	5.94E-6	1.18E-5	2.24E-5
Re-clusterings:	27	56	81
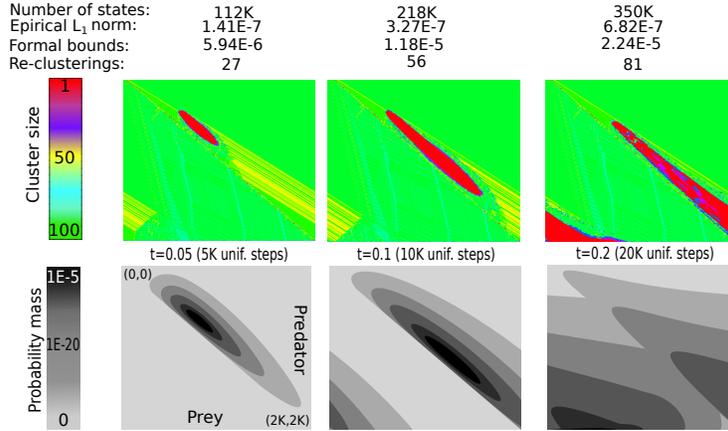
Fig. 1. Transient analysis of the Lotka-Volterra model using history-based adaptive aggregation.

We run all experiments on a MackBook Air™ with 1.8GHz Intel Core i5 and 4 GB 1600 MHz RAM. As expected, for comparable state space reductions (value *avg*), FAU can be faster but in the same order of magnitude as our prototype, due to the overhead of clustering and adaptive uniformisation not being fully integrated in our implementation.

Recall that FAU eliminates states with incoming probability lower than a defined threshold, and as such leads to an under-approximation of the concrete probability distribution with no tailored error bounds: all we can say is that, point-wise, the concrete transient probability distribution resides between this under-approximation and a value obtained by adding the total probability lost, and similarly for the invariance likelihood.

Two-component signalling pathway [7] has analysed the robustness of the output signal of an input-output signal response mechanisms introduced in [28]. It is a two-component signalling pathway including the histidine kinase H, the response regulator R, and their phosphorylated forms (Hp and Rp). In order to ensure a feasible analysis, [7] has limited the state space by bounding the total populations over the intervals $25 \leq H + Hp \leq 35$ and $25 \leq R + Rp \leq 35$ (dimensionless quantities). Since this truncation has a significant impact on the distribution of variable Rp (representing the output signal), in this work we consider less conservative (but computationally more expensive) bounds and employ the adaptive aggregation scheme, which allows for a reduction in the size of the model while quantifying the precision of approximation by means of the derived error bounds.

We first evaluate the adaptive aggregation scheme over the verification of an invariant property with associated small likelihood: in this scenario dynamic truncation techniques such as FAU provide insufficient approximation precision. We compute the probability that the population of Rp stays below the level 15 for $t = 5$ seconds (a relevant time window due to the fast-scale phosphorylation). The results for the new, less restrictive population bounds [5, 55] are reported in Figure 2. We present empirical satisfaction probabilities ("Empirical") and their formal bounds ("Bound") computed using Proposition 3 for the adaptive aggregation scheme, and lower bounds and probability lost for the FAU algorithm. For both schemes we report the obtained state-space

size *avg*. We can observe a clear relationship between the state-space reduction and the precision of the analysis. For adaptive aggregations, the parametrisation of each strategy is denoted by an index (1, 2, 3) representing the thresholds affecting the precision. Note that the parameterization for the history-based aggregation, in contrast to the local strategy, allows us to obtain the user-defined precision (e.g. in this experiment for the history-dependent strategy index 1 denotes a restriction of the bounds to 5E-11, whereas 2 to 5E-12, and 3 to 5E-14), since the aggregation employs exactly the errors. The results also demonstrate that the history-based strategy significantly outperforms the local strategy in all four key performance metrics.

Adaptive aggregation				
Strategy	*avg*	Empirical	Bound	Re-clust.
Local 1	62K	9.55E-12	2.34E-8	65
Local 2	93K	6.32E-13	4.43E-10	81
Local 3	115K	4.54E-14	2.39E-11	97
History 1	54K	5.08E-16	4.68E-11	37
History 2	66K	4.71E-16	4.60E-12	19
History 3	90K	2.20E-16	4.26E-14	20

Fast adaptive uniformisation			
Threshold	*avg*	Lower	Prob. lost
1E-10	15K	0.0	2.68E-5
1E-12	25K	0.0	1.98E-6
1E-15	44K	0.0	1.20E-6
1E-20	91K	0.0	1.00E-6
1E-25	160K	2.12E-17	1.80E-6
1E-30	242K	2.15E-17	1.94E-6

Fig. 2. Statistics for the invariant property. Population bounds [5,55]: 1.2M states (less than those in Fig. 3 due to the property of interest), 16489 steps. The satisfaction probability of the property for the concrete model is equal to 2.15E-17.

Since the invariant property is associated with a small probability, we require accurate error bounds. The data in Figure 2 shows that, for upper bounds of the adaptive scheme that are at least 5 orders of magnitude better compared to those from FAU, the adaptive aggregation method provides more than a twenty-fold reduction with respect to the size of the concrete model, and about a three-fold improvement with respect to the compression obtained via FAU. The results also demonstrate that different parameterisations of the aggregation strategy allow us to control the bounds, and via the bounds also to improve the empirical results (which confirms the usefulness of the derived bounds). However, decreasing the truncation threshold of FAU only improves the lower bounds (from 0.0 to 2.15E-17), but the probability lost is not considerably improved (it is even slightly worse for the very small thresholds, probably due to rounding errors). Notice that, whilst the global errors are still much more conservative, FAU provides better state-space reduction when a lower bound around 2.15E-17 (which is very close to the true probability) is required for the adaptive scheme.

Next, we employ this example to compare the computation of the L_1 norm of the probability distribution at time $t = 5$ seconds. The table in Fig. 3 depicts the results for the L_1 norm over the whole state space, whereas the table in Fig. 4 depicts the results for the L_1 norm over a certain subset of interest. The formal bounds for the adaptive scheme (column "Empirical" in Fig. 3) have been computed using Proposition 2, whilst the corresponding bounds for Fig. 4 (middle part) have been obtained as the sum of the point-wise errors, defined in Proposition 1, over the subset of interest. The upper part of the tables corresponds to the population bounds [25, 35] (as in [7]), whereas the lower part to the less restrictive bounds [5, 55]. Compared to the local strategy, the history-

Population	Adaptive aggregation				Fast adaptive uniformisation		
	Strategy	*avg*	Empirical	Bound	Re-clust.	Threshold	*avg*	Prob. lost
[25, 35]	History 1	70K	1.83E-4	2.56E-2	16	1E-10	72K	1.26E-3
	History 2	88K	2.69E-6	2.95E-4	17	1E-14	105K	1.98E-6
[5, 55]	History 1	453K	2.54E-4	4.73E-2	16	1E-10	132K	1.65E-3
	History 2	515K	4.16E-6	5.31E-4	24	1E-18	493K	1.96E-6

Fig. 3. Statistics for the L_1 norm of the error. Population bounds [25,35]: 116K states, 6924 uniformisation steps. Population bounds [5,55]: 2.5M states, 16489 uniformisation steps.

Pop.	Adaptive aggregation				Fast adaptive uniformisation		
	Strategy	*avg*	Empirical	Bound	Re-clust.	Threshold	*avg*	Empirical	Bound
[25, 35]	Hist. 1	71K	2.92E-8	4.89E-4	18	1E-12	93K	1.77E-9	2.18E-1
	Hist. 2	86K	5.80E-10	3.83E-6	8	1E-14	105K	2.04E-11	2.77E-2
[5, 55]	Hist. 1	354K	2.79E-6	1.75E-3	13	1E-16	388K	7.12E-13	6.02E-1
	Hist. 2	430K	1.10E-8	1.67E-5	19	1E-20	597K	2.68E-14	5.71E-1

Fig. 4. Statistics for the L_1 norm of the error computed over a set, characterised by the population of at least one species that is equal to 0. Population bounds [25,35]: 116K states, 6924 uniformisation steps - the set has 14K states and the probability distribution within the set at time $t = 5$ is equal to 1.31E-8. Population bounds [5,55]: 2.5M states, 16489 uniformisation steps - the set has 307K states and the probability distribution within the set at time $t = 5$ is equal to 1.36E-8.

based aggregation again provides better performance, namely it requires significantly (up to ten-times) smaller numbers of re-clusterings ("Re-clust."): we thus present the results only for the history-dependent strategy. We ensure the comparability of the two outcomes by empirically selecting the threshold for FAU to obtain a truncated model of size (*avg*) similar to that resulting from our technique. Note that, in the case of the L_1 norm over the state space, the probability lost reported by FAU provides the safe bound on the L_1 norm and is equal to the empirical error between the concrete and truncated probability distributions. However, in the case of the L_1 norm over a general subset of the state space the probability lost has to be multiplied by the cardinality of the subset to obtain the correct formal bounds. Such bounds are reported in Fig. 4 (right part) as "Bound," whereas the empirical error between the distribution over the subset is depicted as "Empirical".

Summarising Fig. 3 and Fig. 4, when requiring a tight bound for the smaller state space (population [25, 35]), either approach does not lead to more than a two-fold reduction in the size of the space. This suggests a limit on the possible state-space reduction resulting from the model dynamics. However, for the larger model (population [5, 55]), up to a seven-fold reduction can be obtained using adaptive aggregation. We can see that FAU outperforms the adaptive aggregation scheme in the case of the L_1 norm error over the whole state space (where it leads to a nineteen-fold reduction) but, in contrast to our approach, is not able to provide useful bounds for a general L_1 norm (especially for the larger model).

Prokaryotic gene expression The second case study deals with a more complex model for prokaryotic gene expression. The chemical reaction model has been introduced in [19] and includes 12 species and 11 reactions. We bound the maximal population

Max pop.	Adaptive aggregation				Fast adaptive uniformisation			
	Strategy	*avg*	Empirical	Bound	Re-clust.	Threshold	*avg*	Empirical	Bound
10	Hist. 1	127K	6.36E-7	1.77E-4	35	1E-12	141K	2.55E-7	1.00E+0
10	Hist. 2	207K	2.94E-9	3.77E-7	37	1E-20	386K	7.24E-9	8.05E-1
20	Hist. 1	287K	7.38E-6	2.83E-4	56	1E-12	176K	1.49E-6	1.00E+0
20	Hist. 2	428K	2.56E-8	4.26E-7	59	1E-20	628K	3.35E-8	1.00E+0

Fig. 5. Statistics for the L_1 norm of the error restricted to a set of interest, a strict subset of the state space. Maximal population 10: 1.2M states, 33162 uniformisation steps - the set has 516K states and the probability distribution within the set at time $t = 1000$ is equal to 5.84E-3. Maximal population 20: 4.4M states, 53988 uniformisation steps - the set has 1.8M states and the probability distribution within the set at time $t = 1000$ is equal to 2.21E-2.

of particular species (left column in Fig. 5) to obtain a finite and tractable state space. We focus our experiments exclusively on the history-dependent aggregation scheme.

In contrast with the previous case study that focused on events with very small likelihood, we now discuss results for events with non-negligible likelihood. Figure 5 reports basic statistics on the computation of the L_1 norm over a certain subset of the state space at time $t = 1000$ seconds. Providing useful error bounds on the L_1 norm (computed from the point-wise errors in Proposition 1), the adaptive aggregation leads to almost a ten-fold state space reduction for the smaller model (1.2M vs 127K) and a fifteen-fold reduction for the larger model (4.4M vs 287K). Due to the large cardinality of the subset of interest, FAU fails to provide any informative formal bounds. Note that in this case study the adaptive aggregation scheme also provides better empirical bounds than FAU.

Finally, we have evaluated both approaches on an invariant property (the population of a species stays below the level 10, for 1000 seconds) with a significant satisfaction probability (more than 15% and 20% on the small and large model, respectively). We observe that this choice is favourable to FAU, since for invariant properties with high likelihood the state space truncated via FAU is aligned with the property of interest, and thus the lost probability mass is slightly smaller than the error introduced by the state-space aggregations. In this scenario FAU yields better reductions than the adaptive aggregation scheme (especially for the larger model), while providing similar error bounds, since it is able to successfully identify the relevant part of the state space. This scenario advantageous to FAU is in contrast to that discussed in Figure 2, as well as to the general case where for an arbitrary model it is not known how the probability mass is distributed in relation to the states satisfying the property of interest.

6   Conclusions

We have proposed a novel adaptive aggregation algorithm for approximating the probability of an event in a Markov chain with rigorous precision guarantees. Our approach provides error bounds that are in general orders of magnitude more accurate compared to those from fast adaptive uniformisation, and significantly decreases the size of models without performance degradation. This has allowed us to efficiently analyse larger and more complex models. Future work will include effective combinations of the adaptive aggregation with robustness analysis and parameter synthesis. We also plan to apply our approach to the verification and performance analysis of complex safety-critical computer systems, where precision guarantees play a key role.

References

1. A. Abate, J.-P. Katoen, J. Lygeros, and M. Prandini. Approximate model checking of stochastic hybrid systems. *European Journal of Control*, 16:624–641, 2010.

2. A. Abate, M. Kwiatkowska, G. Norman, and D. Parker. Probabilistic model checking of labelled Markov processes via finite approximate bisimulations. In *Horizons of the Mind. A Tribute to Prakash Panangaden*, volume 8464 of *LNCS*, pages 40–58. Springer, 2014.

3. A. Angius, A. Horvath, and V. Wolf. Quasi product form approximation for Markov models of reaction networks. In *Transactions on Computational Systems Biology (TCSB) XIV*, volume 7625 of *LNCS*, pages 26–52. Springer, 2012.

4. C. Baier and J.-P. Katoen. *Principles of model checking*. The MIT Press, 2008.

5. L. Bortolussi and J. Hillston. Fluid model checking. In *Concurrency Theory (CONCUR)*, volume 7454 of *LNCS*, pages 333–347. Springer, 2012.

6. P. Buchholz. Exact performance equivalence: An equivalence relation for stochastic automata. *Theor. Comput. Sci.*, 215(1-2):263–287, 1999.

7. M. Česka, D. Šafránek, S. Dražan, and L. Brim. Robustness analysis of stochastic biochemical systems. *PloS one*, 9(4):e94553, 2014.

8. T. Chen and S. Kiefer. On the total variation distance of labelled Markov chains. In *Computer Science Logic (CSL) and Logic in Computer Science (LICS)*, 2014.

9. F. Dannenberg, E. M. Hahn, and M. Kwiatkowska. Computing cumulative rewards using fast adaptive uniformisation. *ACM Transactions on Modeling and Computer Simulation, Special Issue on Computational Methods in Systems Biology (CMSB)*, 2014. To appear.

10. J. Desharnais, F. Laviolette, and M. Tracol. Approximate analysis of probabilistic processes: logic, simulation and games. In *Quantitative Evaluation of SysTems (QEST)*, pages 264–273, 2008.

11. A. D'Innocenzo, A. Abate, and J.-P. Katoen. Robust PCTL model checking. In *Hybrid Systems: Computation and Control (HSCC)*, pages 275–285. ACM, 2012.

12. S. Engblom. Computing the moments of high dimensional solutions of the master equation. *Applied Mathematics and Computation*, 180(2):498 – 515, 2006.

13. L. Ferm and P. Lötstedt. Adaptive solution of the master equation in low dimensions. *Applied Numerical Mathematics*, 59(1):187–204, 2009.

14. B. L. Fox and P. W. Glynn. Computing Poisson Probabilities. *Communications of the ACM*, 31(4):440–445, 1988.

15. D. T. Gillespie. Exact Stochastic Simulation of Coupled Chemical Reactions. *Journal of Physical Chemistry*, 81(25):2340–2381, 1977.

16. J. Hasenauer, V. Wolf, A. Kazeroonian, and F. Theis. Method of conditional moments (mcm) for the chemical master equation. *Journal of Mathematical Biology*, pages 1–49, 2013.

17. M. Hegland, C. Burden, L. Santoso, S. MacNamara, and H. Booth. A solver for the stochastic master equation applied to gene regulatory networks. *Journal of computational and applied mathematics*, 205(2):708–724, 2007.

18. T. A. Henzinger, M. Mateescu, and V. Wolf. Sliding Window Abstraction for Infinite Markov Chains. In *Computer Aided Verification (CAV)*, volume 5643 of *LNCS*, pages 337–352. Springer, 2009.

19. A. M. Kierzek, J. Zaim, and P. Zielenkiewicz. The effect of transcription and translation initiation frequencies on the stochastic fluctuations in prokaryotic gene expression. *Journal of Biological Chemistry*, 276(11):8165–8172, 2001.

20. M. Kwiatkowska, G. Norman, and D. Parker. PRISM 4.0: Verification of Probabilistic Real-time Systems. In *Computer Aided Verification (CAV)*, volume 6806 of *LNCS*, pages 585–591. Springer, 2011.

21. K. G. Larsen and A. Skou. Bisimulation through probabilistic testing. *Information and Computation*, 94(1):1 – 28, 1991.
22. C. Madsen, C. Myers, N. Roehner, C. Winstead, and Z. Zhang. Utilizing stochastic model checking to analyze genetic circuits. In *Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pages 379–386. IEEE Computer Society, 2012.
23. M. Mateescu, V. Wolf, F. Didier, and T. A. Henzinger. Fast Adaptive Uniformization of the Chemical Master Equation. *IET Systems Biology*, 4(6):441–452, 2010.
24. B. Munsky and M. Khammash. The finite state projection algorithm for the solution of the chemical master equation. *The Journal of chemical physics*, 124:044104, 2006.
25. R. Sidje and W. Stewart. A numerical study of large sparse matrix exponentials arising in Markov chains. *Computational statistics & data analysis*, 29(3):345–368, 1999.
26. S. E. Z. Soudjani and A. Abate. Adaptive and sequential gridding procedures for the abstraction and verification of stochastic processes. *SIAM Journal on Applied Dynamical Systems*, 12(2):921–956, 2013.
27. S. E. Z. Soudjani and A. Abate. Precise approximations of the probability distribution of a Markov process in time: an application to probabilistic invariance. In *Tools and Algorithms for the Construction and Analysis of Systems (TACAS)*, pages 547–561. Springer, 2014.
28. R. Steuer, S. Waldherr, V. Sourjik, and M. Kollmann. Robust signal processing in living cells. *PLoS computational biology*, 7(11):e1002218, 2011.
29. I. Tkachev and A. Abate. On approximation metrics for linear temporal model-checking of stochastic systems. In *Hybrid Systems: Computation and Control (HSCC)*, pages 193–202. ACM, 2014.
30. A. P. van Moorsel and W. H. Sanders. Adaptive uniformization. *Stochastic Models*, 10(3):619–647, 1994.
31. J. Zhang, L. T. Watson, and Y. Cao. Adaptive aggregation method for the chemical master equation. *International Journal of Computational Biology and Drug Design*, 2(2):134–148, 2009.