A Stochastic Hybrid Approximation for Chemical Kinetics Based on the Linear Noise Approximation

Luca Cardelli[1,2], Marta Kwiatkowska[2], and Luca Laurenti[2]

[1] Microsoft Research
[2] Department of Computer Science, University of Oxford

Abstract. The Linear Noise Approximation (LNA) is a continuous approximation of the CME, which improves scalability and is accurate for those reactions satisfying the leap conditions. We formulate a novel stochastic hybrid approximation method for chemical reaction networks based on adaptive partitioning of the species and reactions according to leap conditions into two classes, one solved numerically via the CME and the other using the LNA. The leap criteria are more general than partitioning based on population thresholds, and the method can be combined with any numerical solution of the CME. We then use the hybrid model to derive a fast approximate model checking algorithm for Stochastic Evolution Logic (SEL). Experimental evaluation on several case studies demonstrates that the techniques are able to provide an accurate stochastic characterisation for a large class of systems, especially those presenting dynamical stiffness, resulting in significant improvement of computation time while still maintaining scalability.

1 Introduction

Biochemical systems are inherently stochastic: the time for the next reaction to occur and which reaction fires next are both random variables. When the reactant molecules are in low number the resulting dynamic behaviour can be highly stochastic and deterministic models are unable to correctly approximate it [23,4]. Thus, an accurate characterisation of stochastic fluctuations in biological systems is essential [30]. It is well known that a biochemical system evolving in a spatially homogeneous environment, at constant volume and temperature, can be described as a continuous-time Markov chain (CTMC) [10] Transient analysis is generally performed through solving the Chemical Master Equation (CME) [30] or with the Stochastic Simulation Algorithm (SSA) [12]. The SSA produces a single realization of the stochastic process, whereas the CME gives the probability distribution of each species over time. The CME can be solved numerically through solving differential equations or methods based on uniformisation, both requiring exploration of the reachable state space and thus infeasible for systems with large or infinite state spaces. On the other hand, the SSA is generally faster, although obtaining good accuracy necessitates potentially large numbers of simulations and can be time consuming.

An alternative is to approximate the CME as a *continuous-state stochastic* process. The *Linear Noise Approximation (LNA)* is a Gaussian process which has been derived as an approximation of the CME [30]. Thus, the LNA is inherently unimodal and not accurate for multimodal dynamics. Its solution involves a number of differential equations that is quadratic in the number of species and independent of the molecular populations. As a consequence, the LNA is generally much more scalable than a discrete stochastic representation. For these reasons, the LNA has recently been used for model checking of large biochemical systems [8,5]. The solution given by the LNA is accurate if conditions on species and reactions known as the *leap conditions* are satisfied, which holds in the limit of high populations, but typically only for a subset of species and reactions (i.e. stiff systems). As a result, a *discrete stochastic* representation is necessary for the remaining species. A natural approach is thus to consider a *stochastic hybrid* semantics that combines a continuous approximation based on the LNA for species respecting the leap conditions and maintains a discrete stochastic representation for the remaining species. Fortunately, for a large class of biological systems the species that respect the leap conditions are in high number [31], which necessitates solving the CME only for a significantly reduced state space.

Contributions. We present a *stochastic hybrid* model for biochemical systems, where a subset of species and reactions is treated with a *continuous* state-space stochastic process, the LNA, while the remaining species are treated as a *discrete* state-space stochastic process. A key advantage is that transient analysis of a discrete stochastic process is needed only for a substantially reduced set of species, ameliorating state-space explosion. The main novelty of our approach is that we partition species and reactions using the leap conditions. This allows us to dynamically and automatically update the partitions, which is necessary since the satisfaction of the leap conditions may change with time. We derive equations for the joint and marginal probability distributions of the partitioned system. Continuous species are modelled as a mixture of Gaussian distributions, enabling us to treat multimodality. We present a numerical method for solving the CME, which adaptively and automatically decides for which species a discrete characterization is needed, and which species can be approximated with the LNA, thus resulting in significant improvement of computation time while still maintaining scalability. We then employ the presented hybrid semantics to build a fast and scalable probabilistic model checking algorithm for Stochastic Evolution Logic (SEL), a temporal logic presented in [8]. We implement the techniques and demonstrate on several case studies their ability to provide an accurate stochastic characterization of systems for which the LNA is imprecise, but full solution of the CME, even using advanced numerical techniques, is not feasible because of scalability issues. We emphasise that our method can be used in conjunction with any existing numerical solution of the CME.

Related Work. The work of Henzinger et al. [18], where a hybrid method is presented with a subset of species treated as a continuous approximation and

the remaining species by solving the CME, differs from ours in at least two key aspects. Firstly, their continuous approximation is *deterministic*, whereas ours is continuous stochastic. Secondly, they partition the species based on a *threshold* on the molecular population, rather than the leap conditions, which may lead to inaccuracies, since the error of the deterministic model depends not only on the molecular population but also on model parameters [10]. Our use of the leap conditions guarantees the accuracy of the stochastic approximation. Thomas et al. [29] develop a conditional LNA method and apply it to gene expression networks. They approximate the probability distribution of gene expression products with the conditional LNA, while still treating promoters with the CME. Our approach is similar in the sense that we also consider the LNA for a subset of the species and a discrete-time Markov process for the remaining ones. However, it is not clear in [29] how to partition the species. Instead, we provide criteria based on the leap conditions to automatically decide for which species the LNA is accurate, and which species instead need a discrete characterization.

In [17], the authors present the method of conditional moments for approximating the moments of the solution of the CME, where small populations are treated via a discrete process and high using approximate moment closure. However, how to automatically partition the species is left as an open problem.

Partitioning of species and reactions of a reaction network for the purpose of speeding up the SSA in multi-scale systems has been explored in [15,26,25]. For instance, Yao et al. introduced the slow-scale stochastic simulation algorithm [6], where they distinguish between fast and slow species. Fast species are then treated assuming they reach equilibrium much faster than the slow ones. Adaptive partitioning of the species has been considered in [19,11]. However, in both cases, the authors consider continuous models that differ from the LNA. In particular, in [11] the authors use a jump diffusion Markov process to approximate the original CTMC and derive error bounds to decide the species partitioning.

2 Background

Chemical Reaction Networks. A *chemical reaction network (CRN)* $C = (\Lambda, R)$ is a pair of finite sets, where Λ is the set of *chemical species*, $|\Lambda|$ denotes its size, and R is a set of reactions. Species in Λ interact according to the reactions in R. A *reaction* $\tau \in R$ is a triple $\tau = (r_\tau, p_\tau, k_\tau)$, where $r_\tau \in \mathbb{N}^{|\Lambda|}$ is the *reactant complex*, $p_\tau \in \mathbb{N}^{|\Lambda|}$ is the *product complex* and $k_\tau \in \mathbb{R}_{>0}$ is the coefficient associated to the rate of the reaction. r_τ and p_τ represent the stoichiometry of reactants and products. Given a reaction $\tau_1 = ([1, 1, 0], [0, 0, 2], k_1)$ we often refer to it as $\tau_1 : \lambda_1 + \lambda_2 \rightarrow^{k_1} 2\lambda_3$. The *state change* associated to a reaction τ is defined by $\upsilon_\tau = p_\tau - r_\tau$. Assuming well mixed environment, constant volume V and temperature, a *configuration* or *state* $x \in \mathbb{N}^{|\Lambda|}$ of the system is given by a vector of the number of molecules of each species. Given a configuration x then $x(\lambda_i)$ represents the number of molecules of λ_i in the configuration and $\frac{x(\lambda_i)}{N}$ is the concentration of λ_i in the same configuration, where $N = V \cdot N_A$ is the volumetric factor, V is the volume and N_A Avogadro's number. The *deterministic*

semantics approximates the concentrations of species over time as the solution $\Phi(t)$ of a set of differential equations of the form:

$$\frac{d\Phi(t)}{dt} = F(\Phi(t)) = \sum_{\tau \in R} \upsilon_\tau \cdot (k_\tau \prod_{i=1}^{|\Lambda|} \Phi_i^{r_{i,\tau}}(t)) \tag{1}$$

where $\Phi_i^{r_{i,\tau}}(t)$ is the ith component of vector $\Phi(t)$ raised to the power of $r_{i,\tau}$, the ith component of vector r_τ. The initial condition is $\Phi(0) = \frac{x_0}{N}$. It is known that Eqn (1) is accurate in the limit of high populations [10].

Stochastic Semantics. The propensity rate α_τ of a reaction τ is a function of the current configuration x of the system such that $\alpha_\tau(x)dt$ is the probability that a reaction event occurs in the next infinitesimal interval dt. We assume mass action kinetics, therefore $\alpha_\tau(x) = k_\tau \frac{\prod_{i=1}^{|\Lambda|} r_{i,\tau}!}{N^{|r_\tau|-1}} \prod_{i=1}^{|\Lambda|} \binom{x(\lambda_i)}{r_{i,\tau}}$, where $r_{i,\tau}$ is the ith component of the vector r_τ, $r_{i,\tau}!$ is its factorial, and $|r_\tau| = \sum_{i=1}^{|\Lambda|} r_{i,\tau}$ [3]. To simplify the notation, N is considered embedded inside the coefficient k_τ for any τ. The stochastic semantics of the CRN $C = (\Lambda, R)$ is represented by a *time-homogeneous continuous-time Markov chain* (CTMC) [10] $(X(t), t \in \mathbb{R}_{\geq 0})$ with state space $S \subseteq \mathbb{N}^{|\Lambda|}$. $X(t)$ is a random vector describing the molecular population of each species at time t. Let $x_0 \in \mathbb{N}^{|\Lambda|}$ be the initial condition of X then $P(X(0) = x_0) = 1$. For $x \in S$, we define $P(x,t) = P(X(t) = x \,|\, X(0) = x_0)$. The transient evolution of X is described by the Chemical Master Equation (CME), a set of differential equations

$$\frac{\mathrm{d}}{\mathrm{d}t}(P(x,t)) = \sum_{\tau \in R} \{\alpha_\tau(x - \upsilon_\tau)P(x - \upsilon_\tau, t) - \alpha_\tau(x)P(x,t)\}. \tag{2}$$

Solving Eqn (2) requires computing the solution of a differential equation for each reachable state. The size of the reachable states depends on the number of species and molecular populations and can be huge or even infinite. As a consequence, solving the CME is generally feasible only for CRNs with very few species and small molecular populations.

Linear Noise Approximation. A promising line of research is to consider continuous state-space approximations of $X(t)$. The *Linear Noise Approximation* (LNA)[30] is a continuous approximation of the CME, which permits a *stochastic* characterization of the evolution of a CRN, while still maintaining scalability comparable to that of deterministic models. The LNA is accurate for processes satisfying the *leap conditions* [31]. Given a CRN $C = (\Lambda, R)$, we say that the Markov process $X(t)$ induced by C satisfies the leap conditions at time t if, for any $\tau \in R$, there exists a finite time interval dt such that:

$$\alpha_\tau(X(t)) \text{ constant in } [t, t+dt] \tag{3}$$

$$\alpha_\tau(X(t)) \cdot dt \gg 1. \tag{4}$$

In [13], Gillespie shows that if these conditions are satisfied then the solution of the CME can be approximated by a *Chemical Langevin Equation (CLE)*. Then, under the assumption that stochastic fluctuations are of the order of $N^{\frac{1}{2}}$ [30,10], we can assume that $X(t)$ admits a solution of the form

$$X(t) = N\Phi(t) + N^{\frac{1}{2}}G(t) \tag{5}$$

where $G(t) = (G_1(t), G_2(t), ..., G_{|\Lambda|})$ is a random vector, independent of N, representing the stochastic fluctuations at time t and $\Phi(t)$ is the solution of Eqn (1). It is possible to show that the probability distribution of $G(t)$ can be modelled by a linear Fokker-Planck equation [31]. For every $t \in \mathbb{R}_{\geq 0}$, $G(t)$ has a multivariate normal distribution whose expected value $E[G(t)]$ and covariance matrix $C[G(t)]$ are the solution of the following differential equations:

$$\frac{\mathrm{d}E[G(t)]}{\mathrm{d}t} = J_F(\Phi(t))E[G(t)] \tag{6}$$

$$\frac{\mathrm{d}C[G(t)]}{\mathrm{d}t} = J_F(\Phi(t))C[G(t)] + C[G(t)]J_F^T(\Phi(t)) + W(\Phi(t)) \tag{7}$$

where $J_F(\Phi(t))$ is the Jacobian of $F(\Phi(t))$, $J_F^T(\Phi(t))$ its transpose, $W(\Phi(t)) = \sum_{\tau \in R} \upsilon_\tau \upsilon_\tau{}^T \alpha_{c,\tau}(\Phi(t))$ and $F_j(\Phi(t))$ the jth component of $F(\Phi(t))$. We assume $X(0) = x_0$ with probability 1; as a consequence $E[G(0)] = 0$ and $C[G(0)] = 0$, which implies $E[G(t)] = 0$ for every t. The following theorem illustrates the nature of the approximation using the LNA.

Theorem 1. *[10] Let $C = (\Lambda, R)$ be a CRN and X the CTMC induced by C. Let $\Phi(t)$ be the solution of Eqn (1) with initial condition $\Phi(0) = \frac{x_0}{N}$ and G be the Gaussian process with expected value and variance given by Eqns (6) and (7). Then, for $t \in \mathbb{R}_{\geq 0}$*

$$N^{\frac{1}{2}}|\frac{X(t)}{N} - \Phi(t)| \Rightarrow_N G(t)$$

In the above $\Rightarrow_N$ indicates convergence in distribution [10]. Theorem 1 shows that $G(t)$ models the stochastic fluctuations around the rate equations and guarantees that the leap conditions are always verified in the limit of high populations. However, they could be satisfied even for relatively small numbers of molecules [31]. To compute the LNA it is necessary to solve $O(|\Lambda|^2)$ first order differential equations, and the complexity is independent of the initial number of molecules of each species. Therefore, one can avoid the exploration of the state space that methods based on uniformization rely upon.

3 Stochastic Hybrid Approximation

The key idea behind our approximation is to partition the species into two classes, those that satisfy the leap conditions, which we approximate by a continuous process using the LNA, and the remaining species, for which we need a discrete model. The stochastic process $X(t)$ induced by the CRN can then be

approximated by a *hybrid* combination of the continuous and discrete processes describing the evolution of the partitions. The set of reactions satisfying the leap conditions may change with time and, as a consequence, the partitions of species and reactions need to adapt with time.
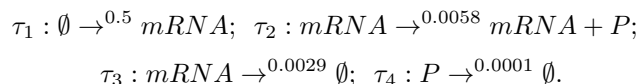
Partitioning of Species and Reactions. Given a CRN $C = (\Lambda, R)$, condition (3) is satisfied for reaction $\tau \in R$ at time t and during the interval dt if $\alpha_\tau(X(t))$ is approximately constant during dt. Reaction $\tau \in R$, at time t, satisfies condition (4) if it fires many times during dt. Given $\sigma_1, \sigma_2 \in \mathbb{R}_{\geq 0}$, it can be equivalently stated that a CRN $C = (\Lambda, R)$ satisfies the leap conditions at time t for an interval dt and reaction $\tau \in R$ if:

$$X_{\lambda_i}(t) \geq \sigma_1 \cdot |v_\tau^{\lambda_i}| \quad \text{for } \lambda_i \text{ such that } v_\tau^{\lambda_i} \neq 0 \text{ and } r_\tau^{\lambda_i} \neq 0 \tag{8}$$

$$\alpha_\tau(X(t)) \geq \sigma_2 \tag{9}$$

where $v_\tau^{\lambda_i}$ represents the state change induced by the occurrence of reaction τ with respect to species λ_i, and $r_\tau^{\lambda_i}$ is the component of the reactant complex relative to species λ_i. A method for choosing $\sigma_1, \sigma_2 \in \mathbb{R}_{\geq 0}$ is given in [26] for SSA (see also below). These criteria induce a partition $R = (R^f, R^s)$ over reactions, where R^f includes reactions for which the leap conditions are satisfied and R^s the remaining reactions, respectively called *continuous* (or fast) reactions and *discrete* (or slow). This induces a partition $\Lambda = (\Lambda^f, \Lambda^s)$ over the species of the CRN, where Λ^f and Λ^s are respectively called *continuous* and *discrete* species. $\lambda \in \Lambda$ is in Λ^f if and only if it is changed by at least one reaction in R^f and it is not changed by reactions in R^s whose propensity is of the same order of magnitude as the reactions in R^f that change it, and otherwise it is in Λ^s. For some systems these criteria may result in species with large populations treated with a discrete stochastic process. This happens for systems where the LNA is not accurate. We illustrate partitioning with the following example.

Example 1. We consider the gene expression model described in [28]. There are two species, $mRNA$ and the protein P, and the following set of reactions

$$\tau_1 : \emptyset \to^{0.5} mRNA; \quad \tau_2 : mRNA \to^{0.0058} mRNA + P;$$

$$\tau_3 : mRNA \to^{0.0029} \emptyset; \quad \tau_4 : P \to^{0.0001} \emptyset.$$

All species are initialized with 0 molecules. We consider $\sigma_1 = 30$ and $\sigma_2 = 0.05$. At time $t = 0$, the initial partition is $\Lambda^f = \{mRNA\}$ and $R^f = \{\tau_1\}$, meaning that the continuous subsystem is given by the only reaction τ_1. In fact, in τ_1 $mRNA$ is not a reagent but only a product. Note that, using a simple threshold on the molecular population of each species to decide if it has a discrete or continuous characterization, as done in [18], would not consider $mRNA$ as a continuous species. After the first molecule of $mRNA$ is produced, the propensity rate of τ_3 increases and its influence needs to be considered. The new species partition becomes $\Lambda^f = \{\}$ and $\Lambda^s = \{mRNA, P\}$. Under our initial conditions,

there exists t' such that $mRNA(t') > 30$ with probability 1. As a consequence, in t' τ_3 is a continuous reaction and the continuous subsystem is:

$$\tau_1 : \emptyset \to^{0.5} mRNA; \ \ \tau_3 : mRNA \to^{0.0029} \emptyset.$$
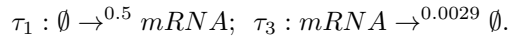
Thus, P is considered a discrete species until both τ_2 and τ_4 become continuous reactions, and thus partitions change over time.

Derivation of the Transient Probability in the Hybrid Model. Based on the partitioning described above, the stochastic process $X(t)$ induced by a CRN can be written as $X(t) = (X^f(t), X^s(t))$, where X^f and X^s respectively describe the evolution of species in Λ^f and species in Λ^s. $X(t)$ is a Markov process, but $X^f(t)$ and $X^s(t)$, if taken separately, are *not* Markovian because they depend on each other. To tackle this issue, following Cao et al. [6], we consider the *virtual* process $\bar{X}^f(t)$ that describes the same species as X^f, but with all the discrete reactions turned off. Therefore, $\bar{X}^f$ is Markovian because it is independent of X^s, and species in Λ^s are now only parameters. Note that $\bar{X}^f$ is only an approximation of the real stochastic process X^f. This approximation is accurate when continuous and discrete species evolve in different time scales. Generally, partitioning using the leap conditions guarantees that. However, it may happen that some reactions satisfy the second leap condition (Eqn 4), but not the first one (Eqn 3). This particular scenario requires attention because these reactions would be classified as discrete, and, in this case, the introduction of the virtual process may introduce some inaccuracies.

Now, we derive equations to study the transient evolution of the continuous and discrete species. Given $x^s \in S^s$ and $x^f \in S^f$, where S^s and S^f are the state spaces of discrete and continuous species, then $P(X^s(t) = x^s, \bar{X}^f(t) = x^f)$, the joint distribution of $X^s(t)$ and $\bar{X}^f(t)$, can be described by the CME (Eqn (2)). However, this would lead to state space explosion. As a consequence, in what follows, we first separate the evolution of continuous and discrete species, and then approximate the continuous subsystem using the LNA. This enables analysis of the transient evolution of the resulting hybrid process.

We denote $P(X^s(t) = x^s, \bar{X}^f(t) = x^f | X^s(0) = x_0^s, \bar{X}^f(0) = x_0^f) = P(x^s, x^f, t)$, $P(X^s(t) = x^s | X^s(0) = x_0^s, \bar{X}^f(0) = x_0^f) = P(x^s, t)$ and $P(\bar{X}^f(t) = x^f | X^s(t) = x^s, \bar{X}^f(0) = x_0^f) = P(x^f | x^s, t)$. Then, as illustrated in [25], by using the axioms of probability we have the following equivalent representation for the CME.

Lemma 1. *Let $x^s \in S^s$ and $x^f \in S^f$. Then, for $t \in \mathbb{R}_{\geq 0}$*

$$\frac{dP(x^f, x^s, t)}{dt} = \frac{dP(x^f | x^s, t)}{dt} P(x^s, t) + P(x^f | x^s, t) \frac{dP(x^s, t)}{dt}$$

So, to solve the CME in this form it is necessary to calculate $P(x^f | x^s, t)$ and $P(x^s, t)$. The first term is Markovian because of the assumption that in the virtual continuous subsystem the continuous species are independent of the discrete

species, which are only parameters. Thus, it can be described by the following master equation for continuous species

$$\frac{dP(x^f|x^s,t)}{dt} = \sum_{\tau \in R^f} \alpha_\tau(x^f - \upsilon_\tau, x^s)P(x^f - \upsilon_\tau|x^s,t) - \alpha_\tau(x^f, x^s)P(x^f|x^s,t) \quad (10)$$

where υ_τ is considered restricted to the components relative to continuous species in $x^f - \upsilon_\tau$. Since the criteria for applicability of the LNA are ensured by partitioning, Eqn (10) can be approximated by the LNA.

On the other hand, $P(x^s,t)$ is not Markovian. However, Proposition 1, whose proof is in the Appendix, guarantees that $P(x^s,t)$ can be derived by solving a set of equations which have the same form as a master equation, and so numerical techniques developed for the CME can still be employed

Proposition 1. *Let $x^s \in S^s$ and $x^f \in S^f$. Then, for $t \in \mathbb{R}_{\geq 0}$ we have*

$$\frac{dP(x^s|t)}{dt} = \sum_{\tau \in R} \beta_\tau(x^s - \upsilon_\tau, t)P(x^s - \upsilon_\tau, t) - \beta_\tau(x^s, t)P(x^s, t) \quad (11)$$

where $\beta_\tau(x^s, t) = \sum_{x^f \in S^f} \alpha_\tau(x^f, x^s)P(x^f|x^s, t)$.

$\beta_\tau(x^s, t)$ is the conditional expectation of the propensity rate of τ at time t given $X^s(t) = x^s$. Reactions of higher order than bi-molecular are not likely [7], and they can always be simulated as a sequence of bi-molecular reactions. As a consequence, we can assume we are limited to at most bi-molecular reactions. Given $\lambda_i^s, \lambda_j^s \in \Lambda^s$ and $\lambda_i^f, \lambda_j^f \in \Lambda^f$, if $\alpha_\tau = k_\tau \cdot \lambda_i^f \cdot \lambda_j^s$ then $\beta_\tau(x^s, t) = k_\tau \cdot E[\bar{X}_{\lambda_i}^f(t)|x^s, t] \cdot x^s(\lambda_j)$. Similarly, if $\alpha_\tau = k_\tau \cdot \lambda_i^f \cdot \lambda_j^f$ then $\beta_\tau(x^s, t) = k_\tau \cdot E[\bar{X}_{\lambda_i}^f(t) \cdot \bar{X}_{\lambda_j}^f(t)|x^s, t]$. If $\alpha_\tau = k_\tau \cdot \lambda_i^s \cdot \lambda_j^s$ then $\beta_\tau(x^s) = k_\tau \cdot x^s(\lambda_i) \cdot x^s(\lambda_j)$. The uni-molecular case follows in a straightforward way. Therefore, to fully characterize $P(x^s, t)$ only the first two moments of the conditional distribution of $\bar{X}^f(t)$ given x^s are needed. In general, this would require solving the entire CME (Eqn (2)). However, thanks to our partitioning criteria, we can safely approximate Eqn (10) by using the LNA and calculating coefficients β using Eqns (6) and (7).

Example 2. Consider the following CRN, taken from [9]:

$$\lambda_z \to^{k_1} \lambda_1; \quad \lambda_z \to^{k_2} \emptyset; \quad \lambda_1 \to^1 \lambda_1 + \lambda_{out}$$
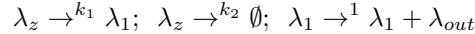
with $k_1, k_2 \in \mathbb{R}_{\geq 0}$ and initial condition x_0 such that $x_0(\lambda_z) = 1$ and $x_0(\lambda_1) = x_0(\lambda_{out}) = 0$. According to the partitioning criteria, for $\sigma_1 > 1$ and $\sigma_2 < \frac{k_1}{k_1 + k_2}$ there exists $t' > 0$ such that for $t > t'$ the set of discrete species is $\Lambda^s = \{\lambda_z, \lambda_1\}$ and the set of continuous species is $\Lambda^f = \{\lambda_{out}\}$ and the partition remains constant over time. A state of the discrete state space is a vector $x^s = (x^s(\lambda_z), x^s(\lambda_1))$. It is easy to verify that the discrete state space S^s is composed of only 3 states: $S^s = \{x_0^s = (1,0), x_1^s = (0,0), x_2^s = (0,1)\}$. According to Eqn (10), and using the law of total probability, the distribution of λ_{out} for $t > t'$ is

given by

$$P(\bar{X}^f_{\lambda_{out}}(t) = k) = P(\bar{X}^f_{\lambda_{out}}(t) = k|x^s_0, t)P(x^s_0, t) +$$
$$P(\bar{X}^f_{\lambda_{out}}(t) = k|x^s_1, t)P(x^s_1, t) + P(\bar{X}^f_{\lambda_{out}}(t) = k|x^s_2, t)P(x^s_2, t)$$

and $P(\bar{X}^f_{\lambda_{out}}(t) = k|x^s_0, t) = P(\bar{X}^f_{\lambda_{out}}(t) = k|x^s_1, t) = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{if } otherwise \end{cases}$. As explained in [9], for $t \to \infty$ we have $P(X^s(t) = x^s_0) = 0$ and $P(X^s(t) = x^s_1) = \frac{k_2}{k_2+k_1}$. As a consequence, our partitioned system correctly predicts that, for $t \to \infty$, λ_{out} has a bimodal distribution that is 0 with probability $\frac{k_2}{k_2+k_1}$.

As shown in Example 2, the distribution of the continuous species can be derived using the law of total probability as $P(x^f, t) = \sum_{x^s \in S^s} P(x^f|x^s, t)P(x^s, t)$. Since each $P(x^f|x^s, t)$ is approximated with the Gaussian distribution given by the LNA, $P(x^f, t)$ is given by a mixture of Gaussian distributions weighted by the probability of being in a particular state of the discrete state space. This enables stochastic characterisation of multimodal distributions for continuous species. Note that the simple LNA, because of its unimodal nature, is unable to represent multimodal behaviours. The following remark shows that, if some assumptions are verified, we can further reduce the computational effort.

Remark 1. Eqn (11) requires solving the LNA once for each $x^s \in S^s$. This can be expensive. However, for a large class of systems, especially those where continuous species have a unimodal distribution, we can consider a reasonable approximation. We can assume $\beta_\tau(x^s, t) \approx \sum_{x^f \in S^f} \alpha_\tau(x^f, E[X^s(t)])$ and $P(x^f, t) = \sum_{x^s \in S^s} P(x^f|x^s, t)P(x^s, t) \approx P(x^f|E[X^s(t)], t)$. So, instead of solving the LNA many times, this requires solving the LNA only once and conditioned on the expectation of the discrete population.

Ensuring Satisfaction of the Leap Conditions. We now explain how to choose constants σ_1 and σ_2 introduced in Eqns (8) and (9). Given a CRN $C = (\Lambda, R)$ and an infinitesimal time interval dt, then $\tau \in R$ satisfies the first leap condition at time t if $\alpha_\tau(X(t))$ is approximately constant during the next dt. This is verified if the relative state change of each reactant species of τ is small enough during dt, that is, if

$$|X_{\lambda_i}(t + dt) - X_{\lambda_i}(t)| \leq max(\epsilon X_{\lambda_i}(t), 1) \text{ for } \lambda_i \in \Lambda \text{ such that } r^{\lambda_i}_\tau \neq 0$$

where $0 \geq \epsilon \geq 1$ is a parameter which quantifies the maximum relative change admitted in reactant species, extensively discussed in [14] for SSA. Rearranging the terms, it is easy to verify that the condition holds if

$$X_{\lambda_i}(t) \geq \frac{|X_{\lambda_i}(t + dt) - X_{\lambda_i}(t)|}{\epsilon} \text{ for } \lambda_i \text{ such that } r^{\lambda_i}_\tau \neq 0 \text{ and } \upsilon^{\lambda_i}_\tau \neq 0.$$

Algorithm 1 Compute Transient Probabilities at Time t_{fin}

Require: A CRN $C = (\Lambda, R)$ with initial condition $x_0 = (x_0^f, x_0^s)$, a finite time interval $[t_0, t_{fin}]$, and parameters for leap conditions σ_1, δ_2.
1: **function** COMPUTEPROB(C, x_0, σ_1, δ_2, $[t_0, t_{fin}]$)
2:      Compute partitions $\Lambda = (\Lambda^f, \Lambda^s)$, $R = (R^f, R^s)$ at time t_0
3:      $(S^s(t_0), X^f(t_0), t) \leftarrow ((x_0^s, 1), x_0^f, t_0)$
4:      **while** $t < t_{fin}$ **do**
5:          Compute Δt and solve discrete master equation for $[t, t + \Delta t]$
6:          **for each** $(x^s, p) \in S^s(t + \Delta t)$ **do**
7:              Solve the LNA to compute $P(X^f(t + \Delta t)|X^s(t) = x^s)$
8:          $t \leftarrow t + \Delta t$
9:          Compute new partitions $\Lambda = (\bar{\Lambda}^f, \bar{\Lambda}^s)$, $R = (\bar{R}^f, \bar{R}^s)$ at time t
10:          **for each** $\lambda_i \in \Lambda$ **do**
11:              **if** $\lambda_i \in \bar{\Lambda}^f \wedge \lambda_i \in \Lambda^s$ **then**
12:                  Move λ_i from $S^s(t)$ to $X^f(t)$
13:              **if** $\lambda_i \in \bar{\Lambda}^s \wedge \lambda_i \in \Lambda^f$ **then**
14:                  Move λ_i from $X^f(t)$ to $S^s(t)$
15:          $(\Lambda^f, \Lambda^s, R^f, R^s) \leftarrow (\bar{\Lambda}^f, \bar{\Lambda}^s, \bar{R}^f, \bar{R}^s)$
16:      $P(X^f(t)) \leftarrow \sum_{(x^s,p) \in S^s(t)} P(X^f(t)|X^s(t) = x^s) \cdot p$
17:      Compute $P(X^s(t))$ by exploration of $S^s(t)$
18:      **return** $(P(X^f(t)), P(X^s(t)))$

Thus, for a given CRN, σ_1 in Eqn (8) quantifies the minimum number of molecules for which we can assume the inequality is satisfied. This is reasonable, as dt is considered to be small, and we assume there are no reactions with unbounded propensity rate. $\tau \in R$ satisfies Eqn (9) if it fires many times during dt, that is, if $\alpha_\tau(X(t)) > \frac{\delta_2}{dt} = \sigma_2$, where δ_2 quantifies the number of times that τ must fire during dt in order to assume the condition satisfied. As a consequence, in order to choose σ_1 and σ_2, we need to tune three parameters: σ_1, δ_2 and dt. Empirical values for σ_1 and δ_2 are given in [26]; dt can be computed as for tau-leaping (see Section 3 of [14]). A possible strategy is to compute dt only once, at time t_0. Then, we can consider dt constant for any $t > t_0$ and make use of Eqns (8),(9). Fixing dt does not affect the correctness of the algorithm, but simply means that, for $t > 0$, there could be a better choice of dt' for which more reactions would be considered continuous.

4   Numerical Implementation

In this section, we present an algorithm to calculate the marginal probability of discrete and continuous species. We first present the general method, where continuous species are modelled as a mixture of Gaussian distributions, and then show how it can be simplified if Remark 1 applies. Algorithm 1 presents the pseudo-code for our routine. In Line 2, we partition species and reactions according to the leap conditions (Eqns (8),(9)). In Line 3, we initialize discrete and

continuous stochastic processes as follows. The discrete process $X^s(t)$ at time t is represented by its state space, $S^s(t)$, given by a set of pairs (x^s, p), where $x^s \in \mathbb{N}^{|\Lambda^s|}$ and p is such that $P(X^s(t) = x^s) = p$. The continuous process $X^f(t)$ at time 0 is equal to x_0^f with probability 1. From Line 4 to 19, the algorithm iteratively updates the partitions. Δt is determined as the integration step of the numerical method used for characterizing discrete species; we use an explicit 4-th order Runge-Kutta algorithm with fixed time step, as in [18]. Alternatively, methods such as uniformisation [22,20] or aggregation-based techniques [1] could also be used. In Line 5, Eqn (11) is solved numerically for the next Δt. In Lines $6 - 7$, for any $(x^s, p) \in S^s(t + \Delta t)$, the algorithm solves the LNA to compute Eqn (10). In Line 9, the partitions are computed at time t according to the leap conditions (Eqns (8), (9)) at that time. In general, the probability mass at time t is distributed over a set of states. In some cases the leap conditions can be checked deterministically based on the expected values $E[X^f(t)]$ and $E[X^s(t)]$. In a more general scenario, it may be necessary to compute the probability that the leap conditions are verified for any $\tau \in R$ and then partition according to these probabilities, which can be approximated as, at time t, we know the approximate solution of the CME [3]. In Lines $11 - 15$, the species are reclassified and the partitions, $S^s(t)$ and $X^f(t)$, are modified accordingly. If λ_i was previously a discrete species and is now assigned to the continuous set, then all states in $S^s(t)$ that are equal except for the number of molecules of λ_i can now be merged. Then, for any state x^s of the updated discrete state space, we compute $P(X_{\lambda_i}^f(t)|X^s(t) = x^s)$, which is Gaussian. In Line 14, for any $(x^s, p) \in S^s(t)$ we discretize the Gaussian distribution $P(X_{\lambda_i}^f(t)|X^s(t) = x^s)$, where $X_{\lambda_i}^f$ is the component of $X^f(t)$ relative to λ_i. Finally, for $t \geq t_{fin}$, in Lines $16 - 17$, the probability distributions of interest are computed.

A Faster Algorithm. If, for a particular CRN, Remark 1 applies then we can assume that $P(X^f(t)) \approx P(X^f(t)|X^s(t) = E[X^s(t)])$. Then we need to compute the LNA only once, and conditioned on the expectation of the discrete stochastic process. The remaining computation can be simplified as well because the virtual continuous process is modelled with a Gaussian distribution and not with a mixture of Gaussians.

Complexity and Error Analysis. The solution of Eqn (11) at time t, using our particular implementation, has a time cost linear in $|S^s(t)|$. We work with the numerical method of [18], which, for each $(x^s, p) \in S^s(t)$, propagates the probability retaining only the x^s such that $P(X^s(t) = x^s) = p > \zeta$. We fix $\zeta = 10^{-14}$. Solving the LNA requires solving a number of differential equations quadratic in the number of continuous species, and independent of the molecular population of such species. In the general case, at time t, we need to solve the LNA during the next Δt a number of times that is of the same order as the dimension of the discrete state space ($O(|S^s(t)||\Lambda^f|)$ differential equations). If Remark 1 is applicable, then the LNA needs to be solved only once.

If all species are partitioned as discrete/continuous, then the solution of Algorithm 1 reduces to that of the CME/LNA. The accuracy depends on the choice of σ_1, σ_2, where it can be shown [14] that, as $\sigma_1, \sigma_2 \to \infty$, then our algorithm guarantees an error equal to the error guaranteed by the numerical method used to solve the discrete master equation. If, instead, both σ_1, σ_2 equal 0, then the error of our hybrid algorithm reduces to the error in computing the LNA, which is model dependent and does not depend only on the molecular counts [10], but also on the validity of assumption (5), which needs to be verified a posteriori [16]. Error bounds would be a viable companion to estimate the committed error, but we are not aware of any explicit formulation of them for the convergence of the LNA. As a result, simulations may be used to validate the results.

5   Model Checking of Stochastic Evolution Logic (SEL)

Employing the hybrid semantics developed here, we present a fast probabilistic model checking algorithms for Stochastic Evolution Logic (SEL) [8]. SEL is a probabilistic logic for analysis of linear combinations of the species of a CRN.

Let $C = (\Lambda, R)$ be a CRN with initial state x_0, then SEL enables evaluation of the probability, variance and expectation of linear combinations of populations of the species of C. The syntax of SEL is given by

$$\eta := P_{\sim p}[B, I]_{[t_1, t_2]} \quad | \quad Q_{\sim v}[B]_{[t_1, t_2]} \quad | \quad \eta_1 \wedge \eta_2 \quad | \quad \eta_1 \vee \eta_2$$

where $Q = \{supV, infV, supE, infE\}$, $\sim = \{<, >\}$, $p \in [0, 1]$, $v \in \mathbb{R}$, $B \in \mathbb{Z}^{|\Lambda|}$, I is a finite set of disjoint intervals and $[t_1, t_2] \subseteq \mathbb{R}_{\geq 0}$. If $t_1 = t_2$ the interval reduces to a singleton.

Formulae η describe global properties of the stochastic evolution of the system. (B, I) specifies a linear combination of the species, where $B \in \mathbb{Z}^{|\Lambda|}$ is a vector defining the linear combination and I represents a set of disjoint closed real intervals. $P_{\sim p}[B, I]_{[t_1, t_2]}$ is the probabilistic operator, which specifies the average value of the probability that the linear combination defined by B falls within the range I over the time interval $[t_1, t_2]$ (we stress that this is not equivalent to reachability). The operators $supE, infE, infV, supV$, see [8], respectively, yield the supremum and infimum of expected value and variance of the random variables associated to B within the specified time interval. The quantitative value associated to a formula can be computed by writing $=?$ instead of $\sim p$ or $\sim v$. For instance, $P_{=?}[B, I]_{[t_1, t_2]}$ gives the probability value associated to the probabilistic property.

Model checking algorithm. Given $Z(t) = B \cdot X(t)$, where B is a linear combination of the species of C, then, according to the semantics of SEL [8], in order to perform model checking, we need to compute $P(Z(t) = z | X(0) = x_0), E[Z(t) | X(0) = x_0]$ and $E[Z(t) \cdot Z(t) | X(0) = x_0]$ (transient probability, expected value and variance of Z), where $z \in \mathbb{Z}$, and $x_0 \in \mathbb{N}^{|\Lambda|}$. In general, this requires solving the CME, which leads to state space explosion ot the LNA,

which is fast but not always accurate. However, we can use our hybrid approximation in order to derive a fast and approximate model checking algorithm of SEL. We approximate Z with Z^h, which is the linear combination of the hybrid approximation of $X = (X^f, X^s)$. The following theorems, whose proofs are in the Appendix, show that model checking SEL just requires computing the hybrid approximation of the CME. In fact, uni-dimensional Gaussian integrals can be computed numerically in constant time. We denote Λ_t^s as the set of discrete species at time t.

Theorem 2. *Assume Λ_t^s is non-empty and S^s is the state space of $X^s(t)$. Then, the stochastic process $Z^h : \Omega \times \mathcal{R}_{\geq 0} \to \mathcal{S}$, with Ω its sample space and $(\mathcal{S}, \mathcal{B})$ a measurable space, is such that for $A \in \mathcal{B}$ and $t \in \mathbb{R}_{\geq 0}$*

$$P(Z^h(t) \in A | X(0) = x_0) = \sum_{x^s \in S^s} P(Z_{x^s}(t) \in A) P(X^s(t) = x_s)$$

where $Z_{x^s}(t)$ is a Gaussian random variable with expected value and variance

$$E[Z_{x^s}(t)] = B \cdot \begin{pmatrix} E[\bar{X}^f(t)] \\ x^s \end{pmatrix} \quad C[Z_{x^s}(t)] = B \cdot \begin{pmatrix} C[\bar{X}^f(t)] & 0 \\ 0 & 0 \end{pmatrix} \cdot B^T$$

where $\bar{X}^f$ is the virtual fast process introduced in Section 3.

Note that if the linear combination, at time t, involves only slow species, then $Z_{x_0}(t)$ is distributed according to a delta-Dirac function. This theorem guarantees that the transient probabilities of Z^h can be computed by solving a set of Guassian integrals, one for each reachable discrete state. The following theorem illustrates that expected value and variance of Z^h can be computed by considering Gaussian properties, even if Z^h is not Gaussian in general.

Theorem 3. *Assume Λ_t^s is non-empty. Then, for $t \in \mathbb{R}_{\geq 0}$*

$$E[Z^h(t) | X(0) = x_0] = \sum_{x^s \in S^s} E[Z_{x^s}(t) \in A] P(X^s(t) = x_s)$$

$$C[Z^h(t) | X(0) = x_0] = \sum_{x^s \in S^s} C[Z_{x^s}(t) \in A] P(X^s(t) = x_s)$$

The basic tools used in the proofs are the law of total expectation and the fact that jointly Gaussian random variables are closed with respect to a linear combination, which is Gaussian [2]. Theorems 2 and 3 assume that, at time t, the set of discrete species is not empty. In fact, if this is the case, all species are treated with the LNA and model checking algorithms for this scenario are given in [8]. We stress that the presented model checking algorithms are accurate only for finite time. In fact, for unbounded time, events that can be neglected in a finite time horizon scenario may fire with probability one. In the next section, SEL is employed in a set of case studies.

6 Experimental Results

We present three case studies showing how our approach significantly improves stochastic analysis of biochemical systems. We implemented Algorithm 1 in Matlab. All the experiments were run on an Intel Dual Core i7 machine with 8 GB of RAM. The first example is a CRN where we need to adaptively partition the species. The second example shows that our hybrid approach can be accurate in cases where the LNA is not, still maintaining comparable time complexity. The third is a system for which advanced numerical techniques for solving the CME such as fast adaptive uniformisation (FAU) [22], as implemented in PRISM [21], fail (out of memory) and using simulations would be too time consuming for comparable accuracy. However, we show that our approach still permits an accurate stochastic characterization.

Gene Expression. We consider the CRN of Example 1. All species in this example follow a unimodal distribution. As a consequence, we employ Remark 1. To ensure a fair comparison, we use the same numerical method for solving the CME and for solving the discrete part of our hybrid model: an explicit 4th order Runge-Kutta algorithm [18]. Even though the stochastic semantics is an infinite CTMC, there are only 2 species in the system with relatively small variance, and thus a numerical solution of the CME is feasible. In Figure 2, in the Appendix, we compare $supE_{=?}[mRNA]_{[T,T]}$ and $supV_{=?}[mRNA]_{[T,T]]}$ for $T \in [0, 200]$, the transient evolution of the expected value and variance of the $mRNA$, as calculated by direct solution of the CME and by our hybrid algorithm. Our algorithm decides to use the LNA for around 70% of the time points. Moreover, we need to adaptively recompute the partitions, as shown in Example 1. In the table below we compare the performance of the same properties for different methods. We consider the following metrics: $||\epsilon||_\infty$ and $||\epsilon||_1$, respectively, average point-wise error and maximum point-wise error of LNA or hybrid approach with respect to the CME solution. $ProbLost$ is the probability lost by the numerical solution of the CME due to the truncation of states with probability mass smaller than 10^{-14}.

Semantics	Time	$||\epsilon||_1$	$||\epsilon||_\infty$	ProbLost
CME	205 sec	-	-	$< 10^{-7}$
Hybrid	35 sec	$< 10^{-7}$	$< 10^{-7}$	-
LNA	5 sec	$9 \cdot 10^{-5}$	0.0112	-

The LNA yields good accuracy. However, our hybrid algorithm achieves accuracy comparable to that for CME and is faster by one order of magnitude.

Bimodal Switch. We consider the CRN presented in Example 2 for $k_1 = 0.7$ and $k_2 = 0.3$. We are interested in analysing the probability distribution of λ_{out} over time, more specifically the SEL property $P_{=?}[\lambda_{out} = K]_{[100,100]}$, for $K \in [0, 200]$. Because of the bimodal nature of such a distribution, Remark 1 is not applicable and the LNA alone is not able to correctly estimate such a distribution. However, our hybrid model, as described in Eqn (2), correctly characterizes the distribution of λ_{out}. Figure 1 compares the distribution of λ_{out}

at time $t = 100$ as estimated by our hybrid approach against the LNA and a full solution of the CME. The following table compares our hybrid approach with
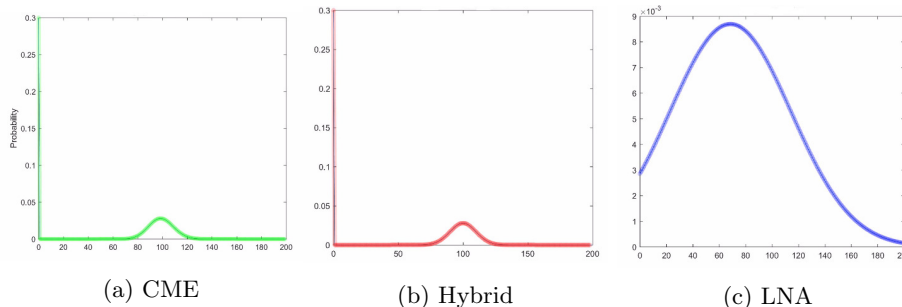

(a) CME (b) Hybrid (c) LNA

Fig. 1: Comparison of the probability distribution of λ_{out} at time $t = 100$, as estimated by a numerical solution of the CME (Fig. 1a), by our hybrid semantics for $\sigma_1 = 2$, $\sigma_2 = 0.5$ (Fig. 1b) and by the LNA (Fig. 1c). Note that in Fig. 1a and 1b there is non-zero probability of having exactly zero molecules.

the other semantics for different values of σ_1 and σ_2. We consider the average point-wise error, $||\epsilon||_1$, and the maximum point-wise error, $||\epsilon||_\infty$, with respect the a numerical solution of the CME, whose error is due to state space truncation ($ProbLost$). For a fair comparison, both for the solution of the master equations of discrete species and for the CME, we use the same numerical method, an explicit 4th order Runge Kutta algorithm with fixed time step [18].

$Semantics$	σ_1	σ_2	$Time$	$||\epsilon||_1$	$||\epsilon||_\infty$	$ProbLost$
CME	-	-	100 sec	-	-	$< 10^{-6}$
LNA	-	-	2.3 sec	0.081	0.2971	-
Hybrid	2	0.5	2.5 sec	$3.284 \cdot 10^{-4}$	0.0024	-
Hybrid	0.5	0.5	2.2 sec	0.081	0.2971	-
Hybrid	2	2	96 sec	$< 10^{-6}$	$< 10^{-6}$	-

For $\sigma_1 > 1$ and $\sigma_2 < 0.7$, the hybrid approach improves the accuracy of the LNA by around two orders of magnitude, while still maintaining comparable execution time. Note that, for this choice of σ_1 and σ_2, the virtual continuous subsystem ignores the delay induced by the firing of the first reaction, which explains why the accuracy of the hybrid method is worse than CME. For $\sigma_2 > 0.7$, all species are considered as discrete and the hybrid approach reduces to the solution of the CME. For $\sigma_1 = \sigma_2 = 0.5$, all species are continuous and the accuracy of the hybrid approach is identical to that of the LNA.

Viral Infection. We consider the intracellular viral infection model proposed in [27]. This model of virus infection is given by the following set of reactions:

$$\tau_1 : DNA + P \rightarrow^{0.00001125} V; \quad \tau_2 : DNA \rightarrow^{0.025} DNA + RNA; \quad \tau_3 : RNA \rightarrow^{0.25}$$

$$\tau_4 : RNA \to^1 RNA + DNA; \quad \tau_5 : RNA \to^{1000} RNA + P; \quad \tau_6 : P \to^{1.9985}$$

The initial condition is $RNA(0) = 1$ and all other species initialized to 0 molecules. We consider $\sigma_1 = 40$ and $\sigma_2 = 20$. This system, although apparently quite small (6 reactions), is very complex to analyse formally or using simulations. This is because it is extremely stiff, with all species presenting high variance and some also high molecular populations. As a consequence, solution of the full CME, even using advanced techniques such as FAU or finite state projection (FSP) [24], is prohibitive due to state-space explosion. For all the properties we consider, FAU is out of memory on our hardware. Because of the stiffness of the system, simulations are time consuming and ensuring good accuracy is not feasible. Our hybrid approach, by considering P as a continuous species for any time instant, enables an effective and efficient stochastic characterization of such a system. Note that, for this system, the LNA is clearly not accurate because of its multimodality.

In Figure 3, in the Appendix, we compare the distribution of the RNA at time $t = 200$ as estimated by our hybrid approach and the distribution of the same species with only the LNA. Results show that the LNA is not able to accurately characterize the distribution of interest, while our hybrid approach correctly predicts multimodality and confirms values obtained by Goutsias in [15] (Figure 5) by using 4000 simulations.

Note that, although the original model is stiff, after species separation the resulting model is much less stiff. This remains true for a large class of systems, and it is a consequence of how we separate the species of a CRN. As a result, for such systems, we need to solve a discrete master equation only for less stiff systems in a reduced state space. As we see in the following table, this results in a marked improvement.

Property (SEL)	Time (Hy)	Time (LNA)	Time (FAU)	RelErr (Hy-LNA)
$P_{=?}[RNA = 0]_{[200,200]}$	4300 sec	28	OutOfMem	0.215
$P_{=?}[RNA = 0]_{[50,50]}$	1500 sec	20	OutOfMem	0.215

$Time(\cdot)$ represents the execution time of different algorithms. $RelErr(Hy\text{-}LNA)$ is the distance between the quantitative value of the property as computed by our hybrid algorithm (and validated by simulations) and by the LNA.

7 Conclusion

We presented a stochastic hybrid approximation of the CME based on automatically partitioning the species and reactions of a CRN according to the leap conditions, and treating the discrete species as a discrete stochastic process, while treating the continuous species as a mixture of Gaussian distributions. The use of the leap conditions justifies the hybrid approximation compared to simple threshold conditions on molecular populations. Our method can be integrated with any numerical method to solve the CME, such as FAU [22], FSP [24] or aggregation based techniques [1]. We demonstrated through case studies that our method is efficient, scales well and can handle multimodality. The algorithm works particularly well for systems where species evolve on different time

scales (i.e. stiff systems), which are common in biology. It also works well when there are no reactions that satisfy the second leap condition, but not the first one. In this case, our hybrid model can introduce some inaccuracies due to the assumptions in partitioning of the species. As future work, we plan to handle this problem by dealing directly with the non-Markovian aspect of the process related to continuous species, without introducing any virtual process. Finally, we plan to implement an algorithm to automatically tune the parameters for species partitioning using stochastic simulations.

References

1. A. Abate, L. Brim, M. Češka, and M. Kwiatkowska. Adaptive aggregation of Markov chains: Quantitative analysis of chemical reaction networks. In *Computer Aided Verification*, pages 195–213. Springer International Publishing, 2015.
2. R. J. Adler. An introduction to continuity, extrema, and related topics for general Gaussian processes. *Lecture Notes-Monograph Series*, 12:i–155, 1990.
3. D. F. Anderson and T. G. Kurtz. *Stochastic analysis of biochemical systems*. Springer.
4. A. Arkin, J. Ross, and H. H. McAdams. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected escherichia coli cells. *Genetics*, 149(4):1633–1648, 1998.
5. L. Bortolussi and R. Lanciani. Model checking Markov population models by central limit approximation. In *Quantitative Evaluation of Systems*, pages 123–138. Springer, 2013.
6. Y. Cao, D. T. Gillespie, and L. R. Petzold. The slow-scale stochastic simulation algorithm. *The Journal of chemical physics*, 122(1):014116, 2005.
7. L. Cardelli. On process rate semantics. *Theoretical Computer Science*, 391(3):190–215, 2008.
8. L. Cardelli, M. Kwiatkowska, and L. Laurenti. Stochastic analysis of chemical reaction networks using linear noise approximation. In *Computational Methods in Systems Biology*, pages 64–76. Springer, 2015.
9. L. Cardelli, M. Kwiatkowska, and L. Laurenti. Programming discrete distributions with chemical reaction networks. In *Proc. 22nd International Conference on DNA Computing and Molecular Programming (DNA22)*, LNCS. Springer, 2016. To appear.
10. S. N. Ethier and T. G. Kurtz. *Markov processes: characterization and convergence*, volume 282. John Wiley & Sons, 2009.
11. A. Ganguly, D. Altintan, and H. Koeppl. Jump-diffusion approximation of stochastic reaction dynamics: Error bounds and algorithms. *Multiscale Modeling & Simulation*, 13(4):1390–1419, 2015.
12. D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361, 1977.
13. D. T. Gillespie. The chemical Langevin equation. *The Journal of Chemical Physics*, 113(1):297–306, 2000.
14. D. T. Gillespie. Simulation methods in systems biology. In *Formal Methods for Computational Systems Biology*, pages 125–167. Springer, 2008.
15. J. Goutsias. Quasiequilibrium approximation of fast reaction kinetics in stochastic biochemical systems. *The Journal of chemical physics*, 122(18):184102, 2005.
16. J. Goutsias and G. Jenkinson. Markovian dynamics on complex reaction networks. *Physics Reports*, 529(2):199–264, 2013.

17. J. Hasenauer, V. Wolf, A. Kazeroonian, and F. Theis. Method of conditional moments (mcm) for the chemical master equation. *Journal of mathematical biology*, 69(3):687–735, 2014.

18. T. A. Henzinger, L. Mikeev, M. Mateescu, and V. Wolf. Hybrid numerical solution of the chemical master equation. In *Proceedings of the 8th International Conference on Computational Methods in Systems Biology*, pages 55–65. ACM, 2010.

19. B. Hepp, A. Gupta, and M. Khammash. Adaptive hybrid simulations for multiscale stochastic reaction networks. *The Journal of chemical physics*, 142(3):034118, 2015.

20. M. Kwiatkowska, G. Norman, and D. Parker. Stochastic model checking. pages 220–270, 2007.

21. M. Kwiatkowska, G. Norman, and D. Parker. PRISM 4.0: Verification of probabilistic real-time systems. In *Computer aided verification*, pages 585–591. Springer, 2011.

22. M. Mateescu, V. Wolf, F. Didier, T. Henzinger, et al. Fast adaptive uniformisation of the chemical master equation. volume 4, pages 441–452. IET, 2010.

23. H. H. McAdams and A. Arkin. Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences*, 94(3):814–819, 1997.

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(4):044104, 2006.

25. C. V. Rao and A. P. Arkin. Stochastic chemical kinetics and the quasi-steady-state assumption: application to the gillespie algorithm. *The Journal of chemical physics*, 118(11):4999–5010, 2003.

26. H. Salis and Y. Kaznessis. Accurate hybrid stochastic simulation of a system of coupled chemical or biochemical reactions. *The Journal of chemical physics*, 122(5):054103, 2005.

27. R. Srivastava, L. You, J. Summers, and J. Yin. Stochastic vs. deterministic modeling of intracellular viral kinetics. *Journal of Theoretical Biology*, 218(3):309–321, 2002.

28. M. Thattai and A. Van Oudenaarden. Intrinsic noise in gene regulatory networks. *Proceedings of the National Academy of Sciences*, 98(15):8614–8619, 2001.

29. P. Thomas, N. Popović, and R. Grima. Phenotypic switching in gene regulatory networks. *Proceedings of the National Academy of Sciences*, 111(19):6994–6999, 2014.

30. N. G. Van Kampen. *Stochastic processes in physics and chemistry*, volume 1. Elsevier, 1992.

31. E. Wallace, D. Gillespie, K. Sanft, and L. Petzold. Linear noise approximation is valid over limited times for any chemical system that is sufficiently large. *IET systems biology*, 6(4):102–115, 2012.

A Proofs

Proposition 1. *Let $x^s \in S^s$ and $x^f \in S^f$. Then, for $t \in \mathbb{R}_{\geq 0}$*

$$\frac{dP(x^s|t)}{dt} = \sum_{\tau \in R} \beta_\tau(x^s - \upsilon_\tau, t)P(x^s - \upsilon_\tau, t) - \beta_\tau(x^s, t)P(x^s, t)$$

where $\beta_\tau(x^s, t) = \sum_{x^f \in S^f} \alpha_\tau(x^f, x^s)P(x^f|x^s, t)$.

Proof. By using the law of total probability we have

$$\frac{dP(x^s|t)}{dt} = \sum_{x^f \in S^f} \frac{dP(x^s, x^f, t)}{dt}$$

Then, using Eqn (2), and rearranging terms we have

$$\sum_{x^f \in S^f} \frac{dP(x^s, x^f, t)}{dt} =$$

$$\sum_{x^f \in S^f} \sum_{\tau \in R^f} \alpha_\tau(x^f - \upsilon_\tau, x^s - \upsilon_\tau) P(x^f - \upsilon_\tau, x^s - \upsilon_\tau, t) - \alpha_\tau(x^f, x^s) P(x^f, x^s, t) =$$

$$\sum_{\tau \in R} \beta_\tau(x^s - \upsilon_\tau, t) P(x^s - \upsilon_\tau, t) - \beta_\tau(x^s, t) P(x^s, t)$$

where $\beta_\tau(x^s, t) = \sum_{x^f \in S^f} \alpha_\tau(x^f, x^s) P(x^f|x^s, t)$, that is, the conditional expectation of the propensity rate of τ at time t given $X^s(t) = x^s$.

Theorem 2. *Assume Λ_t^s is non-empty and S^s is the state space of $X^s(t)$. Then, the stochastic process $Z^h : \Omega \times \mathcal{R}_{\geq 0} \to \mathcal{S}$, with Ω its sample space and $(\mathcal{S}, \mathcal{B})$ a measurable space, is such that for $A \in \mathcal{B}$ and $t \in \mathbb{R}_{\geq 0}$*

$$P(Z^h(t) \in A | X(0) = x_0) = \sum_{x^s \in S^s} P(Z_{x^s}(t) \in A) P(X^s(t) = x_s)$$

where $Z_{x^s}(t)$ is a Gaussian random variable with expected value and variance

$$E[Z_{x^s}(t)] = B \cdot \begin{pmatrix} E[\bar{X}^f(t)] \\ x^s \end{pmatrix} \quad C[Z_{x^s}(t)] = B \cdot \begin{pmatrix} C[\bar{X}^f(t)] & 0 \\ 0 & 0 \end{pmatrix} \cdot B^T$$

where $\bar{X}^f$ is the virtual fast process introduced in Section 3.

Proof. By the law of total probability we have

$$P(Z(t) \in A | X(0) = x_0) = \sum_{x^s \in S^s} P(Z(t) \in A | X^s(t) = x^s, X(0) = x_0) P(X^s(t) = x^s | X(0) = x_0).$$

By application of the LNA it follows that $X^f(t)$ conditioned on the event $X^s(t) = x^s$ is a Gaussian random variable with expected value and variance

$$E[X^f(t)|X^s(t) = x^s] = \begin{pmatrix} E[\bar{X}^f(t)] \\ x^s \end{pmatrix}$$

and covariance matrix

$$C[X^f(t)|X^s(t) = x^s] = \begin{pmatrix} C[\bar{X}^f(t)] & 0 \\ 0 & 0 \end{pmatrix}$$

Given a multidimensional Gaussian distribution, each linear combination of its components is still Gaussian. As a consequence, $E[Z^h(t)|X^s(t) = x^s] = B \cdot E[X^f(t)|X^s(t) = x^s]$ and $C[Z^h(t)|X^s(t) = x^s] = B \cdot C[X^f(t)|X^s(t) = x^s] \cdot B^T$.

Theorem 3. *Assume Λ_t^s is non-empty. Then, for $t \in \mathbb{R}_{\geq 0}$*

$$E[Z^h(t)|X(0) = x_0] = \sum_{x^s \in S^s} E[Z_{x^s}(t) \in A]P(X^s(t) = x_s)$$

$$C[Z^h(t)|X(0) = x_0] = \sum_{x^s \in S^s} C[Z_{x^s}(t) \in A]P(X^s(t) = x_s)$$

Proof. The proof follows from the application of the law of total expectation for random variables with mutually exclusive and exhaustive events.
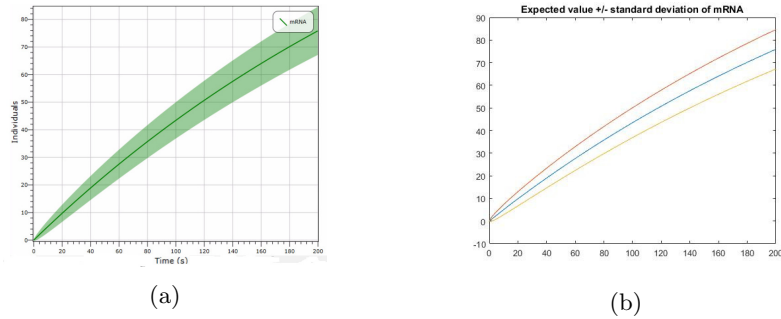
B Figures

(a)

(b)

Fig. 2: Comparison of expected value and variance of $mRNA$ in Example 2 in interval $[0, 200]$ as calculated by direct solution of the CME (Fig. 2a) and by our algorithm (Fig. 2b).
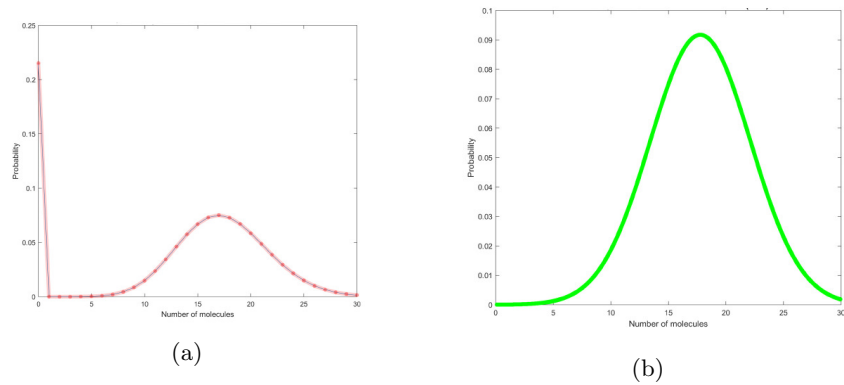
(a)

(b)

Fig. 3: Comparison of the probability distribution of RNA at time $t = 200$ as calculated by numerical hybrid algorithm (Fig. 3a) and by the LNA (Fig. 3b).