DelaySSAToolkit.jl

Citation for published version:
https://doi.org/10.1093/bioinformatics/btac472

Digital Object Identifier (DOI):
10.1093/bioinformatics/btac472

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Bioinformatics

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
DelaySSAToolkit.jl: stochastic simulation of reaction systems with time delays in Julia

Xiaoming Fu et al.

MOE Key Laboratory of Smart Manufacturing in Energy Chemical Process, East China University of Science and Technology, China

School of Biological Sciences, the University of Edinburgh, United Kingdom

Contact: xmfu@ecust.edu.cn, zcao@ecust.edu.cn or ramon.grima@ed.ac.uk

The stochastic simulation algorithm (SSA) is widely used to simulate the time-dependent trajectories for complex systems with Markovian dynamics. A major assumption behind these models is the memoryless hypothesis, i.e., the stochastic dynamics of the reactants is only influenced by the current state of the system, which implies that the waiting times between successive reaction events follow an exponential distribution.

Many common reaction systems encapsulate multiple intermediate reaction steps involving a large number of interacting species. This leads to computationally expensive stochastic simulations using the SSA, which seriously limits the exploration of dynamics across large portions of parameter space. A viable alternative is the use of reduced models; while
rigorous analytical reductions are possible if timescale separation exists \citep{mastny2007two}, more often than not it is more practical to replace multiple intermediate reactions by a single delayed reaction. This leads to a reduced model with non-Markovian dynamics, i.e. where the waiting times are non-exponential, which cannot be simulated by the conventional SSA.

Several exact simulation methods for reaction systems with delays were proposed \citep{Barrio2005,Xiaodong2007}. Some of these methods are available via software tools \citep{Maarleveld2013,Barbuti2009}. However, these packages offer limited functionality (implementing only few delay SSA algorithms and restricted to fixed delays, or mass action reactions) thus limiting the potential user base.

We present DelaySSAToolkit.jl, the first Julia package for modelling stochastic reaction systems with time delay. These delays can be either fixed or vary in time, i.e. each time a delay reaction fires, the delay is chosen from some user-defined probability distribution. It is also possible to model the case of heterogeneous time delays where there is a population of identical reactant systems, each of which has a reaction with a different but fixed time delay. As well, the toolkit can integrate simultaneous delay reactions and cascades of delay reactions using a variety of delay stochastic simulation algorithms. A comprehensive review of delay SSA theory and tutorials on using the software can be found at https://palmtree2013.github.io/DelaySSAToolkit.jl/dev.

\section{Materials and methods}

In what follows, we discuss the main implementation details of DelaySSAToolkit.jl and its integration within the broader Julia package ecosystem.

enabling a streamlined modelling workflow in which we can easily define a Markovian reaction network, extend it to a delay system and solve the resulting system using a chosen delay stochastic simulation algorithm either on a trajectory-level or on a distribution-level via ensemble parallel computing.

In Figure \ref{Fig1}, we summarize the workflow and showcase the software applied to an epidemic model, a birth-death stochastic model with heterogeneous delays \citep{Cortez2021} and an auto-regulatory network of oscillatory gene expression with delays \citep{Jiang2021}.

\subsection{Model definition}

The definition of a delay reaction system includes two parts, namely a Markovian part where the reactions affect the state of reactants instantaneously, and a non-Markovian part where the change of the state of reactants happens after a certain time delay $\tau$ (Figure \ref{Fig1}a). The Markovian part of reaction networks can be constructed by specifying the reaction stoichiometry and the propensities using the Julia symbolic-numeric modelling framework embodied within Catalyst.jl (https://github.com/SciML/Catalyst.jl) and ModelingToolkit.jl \citep{ModelingToolkit}. Another way is via DiffEqJump.jl \citep{Rackauckas2017} by manually defining the stoichiometry and propensity of each reaction using a lower level programming interface.

DelaySSAToolkit.jl extends the Markovian models defined above to a delay system by specifying the causal relationships between the instant reactions and delayed reactions triggered by them. As such, DelaySSAToolkit.jl can handle a non-Markovian system containing any number of species and reactions with any type of smooth propensity functions with a wide range of delay types.
The problem definition, time stepping and algorithms implementation of DelaySSAToolkit.jl are provided by DiffEqJump.jl and the solution handling uses DifferentialEquations.jl. It currently supports four delay stochastic simulation algorithms: delay rejection method, delay direct method, delay modified next reaction method and delay direct method with composition and rejection. We note that the delay rejection and delay direct methods can offer better performance for small reaction networks, while the other two are preferable for systems with significant number of reactions. The computational efficiency of delay modified next reaction method and delay direct method with composition and rejection is improved for large-scale networks using the dependency graph and priority queue described in \citep{Gibson2000}. Ensemble simulation is made easy using high-performance multithreading/multiprocessing parallel computing interface implemented in DifferentialEquations.jl \citep{Rackauckas2017}, which also provides a number of numerical analysis and parameter estimation tools enabling even further study of the resulting delay system.

The resultant delay SSA solution can be given at user-specified time points or else at the exact event time points. The data structure of each solution is composed of three parts: the time points, the state of the reactants and the state of the delay channels at the corresponding time points. The recorded state of delay channels can be particularly useful in some cases, e.g. when modelling gene transcription, the delay reaction models the time between initiation and termination of transcription, and from the state of the delay channel one can reconstruct the positions of RNA polymerases on the gene \citep{fu2021accurate}.
in the Julia ecosystem\citep{Rackauckas2017,ModelingToolkit,Sukys2021,roesch2021julia}, DelaySSAToolkit.jl makes the stochastic modelling of biochemical reaction kinetics efficient and accessible for Julia newcomers and experts alike.\vspace*{-16pt}

\section*{Funding}
This work was supported by NSFC 62073137 to X.F. and Z.C., and a fellowship of the CPSF 2021M071202 and Shanghai Sailing Program 22YP1410700 to X.F.\vspace*{-16pt}

\section*{Acknowledgements}
We thank Augustinas Sukys for useful discussions, feedback and independent testing of the Julia package. \vspace*{-16pt}

\%noindent\textit{Conflict of Interest}: none declared.\vspace*{-12pt}

\bibliographystyle{natbib}
\bibliography{document.bib}
\end{document}