SAND2019-3529PE

Explicit Time Integration of Stiff Chemical Langevin Equations using Computational Singular Perturbation

Habib N. Najm,[1] Xiaoying Han,[2] Mauro Valorani[3]

[1] Sandia National Laboratories, Livermore, CA

[2] Auburn University, Auburn, AL

[3] Sapienza University, Rome, Italy

MIT Seminar
Cambridge, MA
Mar 20-22, 2019

oratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, neywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract

Acknowledgement

This work was supported by:

- DOE Office of Basic Energy Sciences, Div. of Chem. Sci., Geosci., & Biosci.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. The views expressed in the article do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

Outline

Motivation

- Fundamentally the chemical master equation (CME) governs the evolution of chemical systems at the smallest scales
  - Discrete Markov system - integer valued molecular counts
- The Chemical Langevin Equation (CLE) is good approximation for the CME when the number of molecules of each species in the given control volume is large enough – continuous Markov system
  - Technically, the approximation relies on conditions on the reaction rates and time step sizes, which are essentially guaranteed in the context of large molecular counts (Gillespie JChemPhys 2000)
- The CLE is relevant when the number of molecules of each species is small enough so that stochastic effects are non-negligible
  - macroscale deterministic models are inadequate
- Relevant applications:
  - biochemistry
  - catalysis

Stochastic Chemical System Formulation – 1

- Consider a chemical system involving
  - N species $\mathcal{S}_1, \cdots, \mathcal{S}_N$
  - R reactions $\mathcal{R}_1, ..., \mathcal{R}_R$.
- Assume: spatially uniform, fixed volume, constant temperature
- $X_i(t)$: number of $\mathcal{S}_i$ molecules at time t, and state vector:

$$\boldsymbol{X}(t) := (X_1(t), \cdots, X_N(t))^T.$$

- Under requisite conditions, the system can be modelled by the chemical Langevin difference equation, where for $i = 1, \cdots, N$:

$$X_i(t+\mathrm{d}t) = X_i(t) + \sum_{j=1}^{R} \nu_{ji} \rho_j(\boldsymbol{X}(t))\mathrm{d}t + \sum_{j=1}^{R} \nu_{ji} \sqrt{\rho_j((\boldsymbol{X}(t))} \mathcal{N}_j(t) \sqrt{\mathrm{d}t}$$

where

- ν_{ji} is the change in X_i caused by one $\mathcal{R}_j$ reaction
- ρ_j is the propensity function for reaction $\mathcal{R}_j$
- $\mathcal{N}_j(t)$ are *iid* standard normal random variables at time t

Stochastic Chemical System Formulation – 2

The above stochastic difference equation implies the equivalent chemical Langevin differential equation

$$\mathrm{d}X_i(t) = \sum_{j=1}^{R} \nu_{ji}\rho_j(\boldsymbol{X}(t))\mathrm{d}t + \sum_{j=1}^{R} \nu_{ji}\sqrt{\rho_j((\boldsymbol{X}(t))}\mathrm{d}W_j(t), \quad i = 1, \cdots, N$$

where $W_j(t)$ are statistically independent Brownian motions.

We can write the CLE, for convenience, as

$$\mathrm{d}\boldsymbol{X}_t = f(\boldsymbol{X}_t)\mathrm{d}t + \sum_{j=1}^{R} g_j(\boldsymbol{X}_t)\mathrm{d}W_j(t)$$

where $\boldsymbol{X}_t := \boldsymbol{X}(t)$ and

$$f(\boldsymbol{X}_t) = (f_1(\boldsymbol{X}_t), \ldots, f_N(\boldsymbol{X}_t))^T, \qquad f_i(\boldsymbol{X}_t) = \sum_{j=1}^{R} \nu_{ji}\rho_j(\boldsymbol{X}(t))$$

$$g_j(\boldsymbol{X}_t) = (g_{j1}(\boldsymbol{X}_t), \ldots, g_{jN}(\boldsymbol{X}_t))^T, \qquad g_{ji}(\boldsymbol{X}_t) = \nu_{ji}\sqrt{\rho_j((\boldsymbol{X}(t))}$$

SDE Time Integration

- The time integration of the CLE can employ a range of available time integration schemes for stochastic differential equations (SDEs)
- Consider the Itô SDE

$$\mathrm{d}X_t = f(X_t)\mathrm{d}t + g(X_t)\mathrm{d}W_t$$

- Time integration

$$X_t = X_{t_0} + \int_{t_0}^{t} f(X_s)\mathrm{d}s + \int_{t_0}^{t} g(X_s)\mathrm{d}W_s$$

- Euler-Maruyama (EM) – explicit, order 1 weak convergence

$$Y_{n+1} = Y_n + f_n h_n + g_n \sqrt{h}_n \mathcal{N}_n, \qquad h_n = t_{n+1} - t_n, \quad Y_0 = X_{t_0}$$

  - EM is the simplest explicit SDE time integration

SDE Stiffness

- An SDE is stiff when it exhibits a large range of time scales
- A chemical system with very slow/fast reactions results in a stiff CLE
- Stiffness results in challenges for **explicit** SDE time integrators
  - Stability requires time steps smaller than the fastest time scale
  - However, for accurate time integration, ideally, the optimal time step choice is dictated by the *active* time scale, *i.e.* the time scale at which the state vector is changing
- One remedy is to use implicit time integration, but can we do better with explicit constructions?
  - This has been done for ODEs using Computational Singular Perturbation (CSP)

    Valorani & Goussis, JCP, 2001
  - We would like to extend this to SDEs, specifically to the CLE

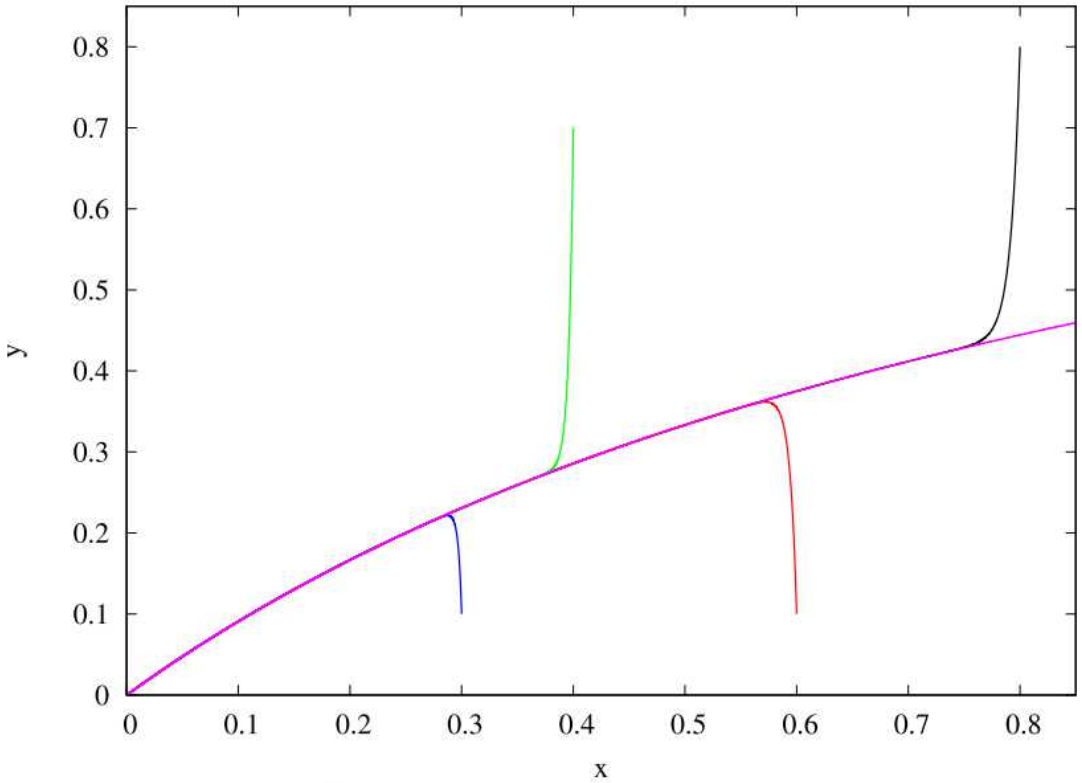
ODE Stiffness

- A macroscale chemical system with very slow & fast reactions results in a stiff ODE system – challenging for **explicit** ODE time integrators
- A stiff ODE system $d\boldsymbol{u}/dt = g(\boldsymbol{u})$, with large negative eigenvalues of the Jacobian $J = \partial g/\partial u$, behaves essentially like a Differential Algebraic Equation (DAE)
  - The system dynamics exhibit attractive slow invariant algebraic manifolds $G(\boldsymbol{u}) = 0$
    - We leave aside for now unstable manifolds, oscillatory and chaotic dynamics
  - From any initial condition, the system state evolves quickly, with the fast time scales, towards the manifold
  - At the manifold, the relevant fast process is "exhausted", inducing the constraint $G(\boldsymbol{u}) = 0$
  - The system evolves slowly, with the slow time scales, along the slow manifold

Davis-Skodje stiff ODE system – example

(Davis & Skodje, J. Chem. Phys., 1999)

$\epsilon = 0.01$



Manifold:

$$y = \frac{x}{1+x}$$

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} &= -x \\ \frac{\mathrm{d}y}{\mathrm{d}t} &= -\frac{1}{\epsilon}y + \frac{1}{\epsilon}\frac{x}{1+x} - \frac{x}{(1+x^2)} \end{aligned}$$

Computational Singular Perturbation (CSP) Basics

Harvey Lam, Dimitris Goussis, 1980s –

- Stiff ODEs are singularly perturbed differential equations
- A decoupled fast-slow stiff system

$$\begin{aligned} \dot{x} &= g_s(x,y) \\ \dot{y} &= \frac{1}{\epsilon} g_r(x,y) \qquad \epsilon \ll 1 \end{aligned}$$

  where x is the slow variable, and y is the fast variable

- Associated DAE system, in the limit $\epsilon \to 0$,

$$\begin{aligned} \dot{x} &= g_s(x,y) \\ 0 &= g_r(x,y) \end{aligned}$$

- This simple distinction of slow/fast variables is not feasible in practical systems
- CSP is useful in decoupling fast and slow processes

CSP Basics – 2

- Consider a general autonomous stiff ODE system

$$\dot{\boldsymbol{x}} = \boldsymbol{g}(\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathbb{R}^N$$

- Introduce:
  - CSP basis vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N$, with $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_N]$
  - CSP row vectors $\boldsymbol{b}^1, \ldots, \boldsymbol{b}^N$, with $\boldsymbol{B}$: $\text{row}_i := \boldsymbol{b}^i$
  - where $\boldsymbol{b}^i \boldsymbol{a}_j = \delta_{ij}$, thus $\boldsymbol{B}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{B} = \boldsymbol{I}$

- Expand the RHS in the CSP basis

$$\boldsymbol{g}(\boldsymbol{x}) = \sum_{i=1}^{N} \boldsymbol{a}_i f^i$$

where

$$f^i(\boldsymbol{x}(t)) = \boldsymbol{b}^i(\boldsymbol{x}(t)) \cdot \boldsymbol{g}(\boldsymbol{x}(t))$$

CSP Basics – 3

Whence

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} f^1 \\ \vdots \\ f^N \end{pmatrix} = \Lambda(\boldsymbol{x}(t))\begin{pmatrix} f^1 \\ \vdots \\ f^N \end{pmatrix}$$

where

$$\Lambda = \begin{bmatrix} (\frac{\mathrm{d}\boldsymbol{b}^1}{\mathrm{d}t} + \boldsymbol{b}^1 J_{\boldsymbol{g}})\boldsymbol{a}_1 & \cdots & (\frac{\mathrm{d}\boldsymbol{b}^1}{\mathrm{d}t} + \boldsymbol{b}^1 J_{\boldsymbol{g}})\boldsymbol{a}_N \\ \vdots & \cdots & \vdots \\ (\frac{\mathrm{d}\boldsymbol{b}^N}{\mathrm{d}t} + \boldsymbol{b}^N J_{\boldsymbol{g}})\boldsymbol{a}_1 & \cdots & (\frac{\mathrm{d}\boldsymbol{b}^N}{\mathrm{d}t} + \boldsymbol{b}^N J_{\boldsymbol{g}})\boldsymbol{a}_N \end{bmatrix} \in \mathbb{R}^{N\times N}$$

and $J_g = \partial \boldsymbol{g}/\partial \boldsymbol{x}$ is the Jacobian of $\boldsymbol{g}$

- The ideal basis decouples fast and slow processes, *i.e.* diagonalizes Λ
  - The eigenvectors of $J_{\boldsymbol{g}}$ provide a good, $\mathcal{O}(\epsilon)$, approximation of the ideal CSP basis
    - This is exact for a linear system, where $\mathrm{d}\boldsymbol{b}^i/\mathrm{d}t \equiv 0, \forall i$
  - A refinement procedure is available to maximize the degree of decoupling starting from any initial guess

CSP Illustration with the Davis-Skodje ODE system

With $\boldsymbol{z} := \begin{bmatrix} y \\ x \end{bmatrix}$ and $\gamma := 1/\epsilon$: $\quad \dot{\boldsymbol{z}} = \boldsymbol{g}(\boldsymbol{z}) = \begin{bmatrix} -\gamma y + \dfrac{\gamma x}{1+x} - \dfrac{x}{(1+x)^2} \\ -x \end{bmatrix}$

Jacobian $\quad \boldsymbol{J} = \begin{bmatrix} -\gamma & \theta \\ 0 & -1 \end{bmatrix}, \quad \theta = \dfrac{\gamma}{(1+x)^2} + \dfrac{x-1}{(1+x)^3}$

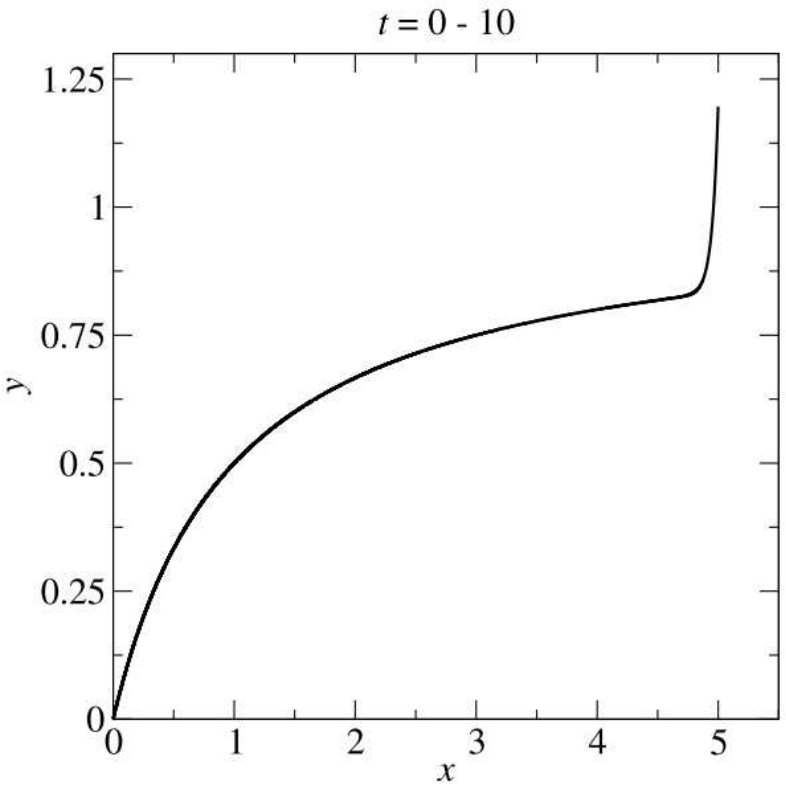
Eigenvalues $\quad \lambda_1 = -\gamma, \quad \lambda_2 = -1$

Basis vectors $\quad \boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \boldsymbol{a}_2 = \begin{bmatrix} \theta \\ \gamma - 1 \end{bmatrix}, \boldsymbol{b}^1 = \left[1, \dfrac{-\theta}{\gamma - 1}\right], \boldsymbol{b}^2 = \left[0, \dfrac{1}{\gamma - 1}\right]$

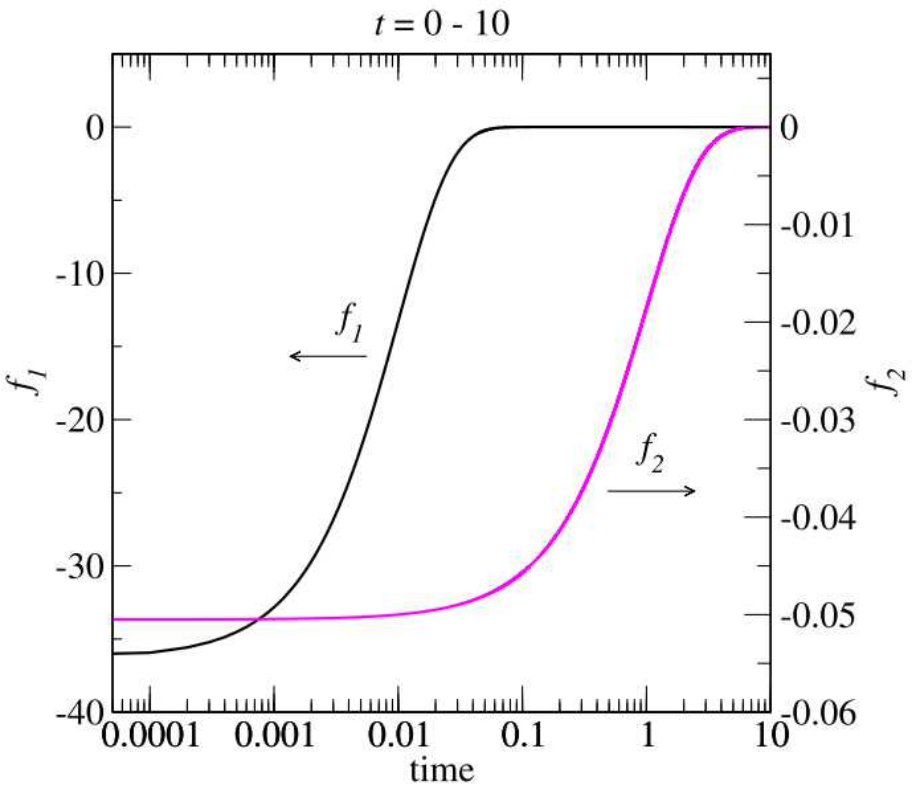
Modes $\quad f^1 = \boldsymbol{b}^1 \cdot \boldsymbol{g} = -\gamma y + \dfrac{\gamma x}{1+x} - \dfrac{x}{(1+x)^2} + \dfrac{x\theta}{\gamma - 1}$

$\qquad\quad f^2 = \boldsymbol{b}^2 \cdot \boldsymbol{g} = \dfrac{-x}{\gamma - 1}$

CSP Illustration with the Davis-Skodje ODE system

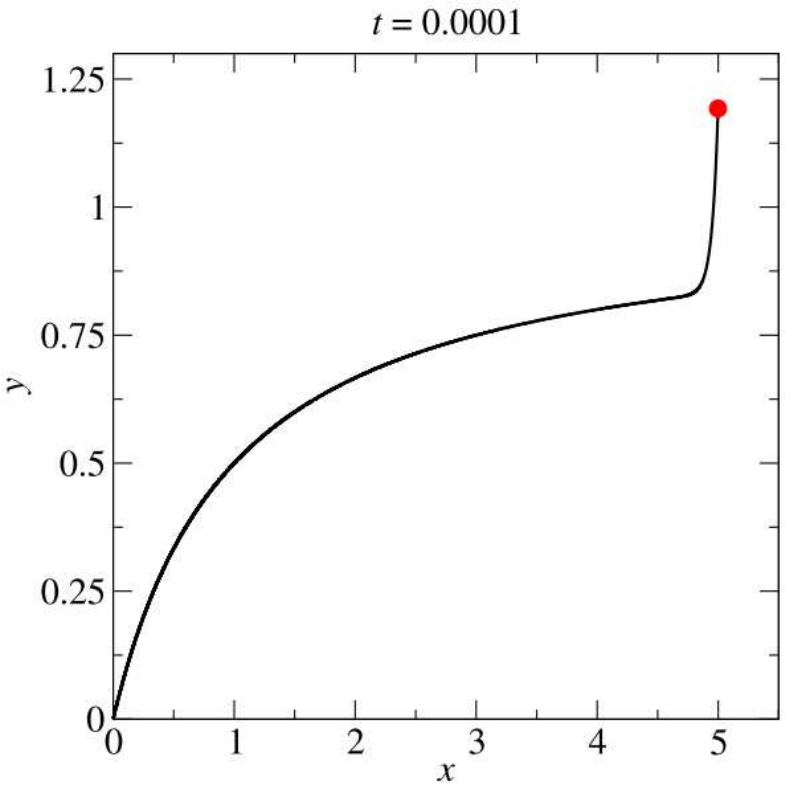


Time evolution of the state vector in the configuration space

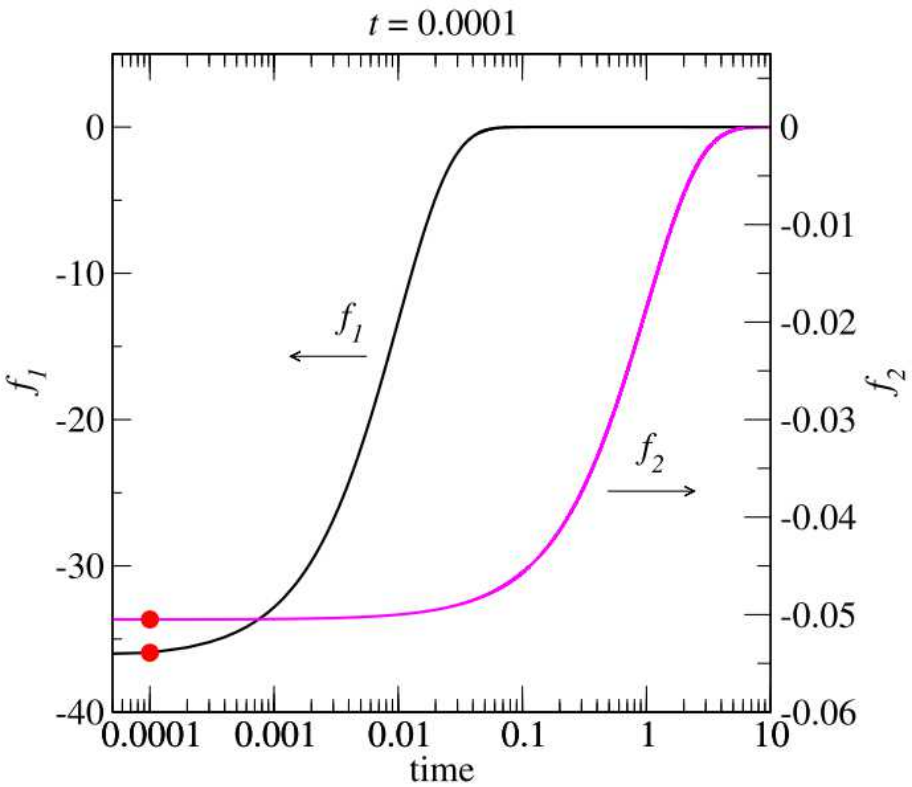


Time evolution of CSP (signed) mode amplitudes $f^i, i = 1, 2$

CSP Illustration with the Davis-Skodje ODE system

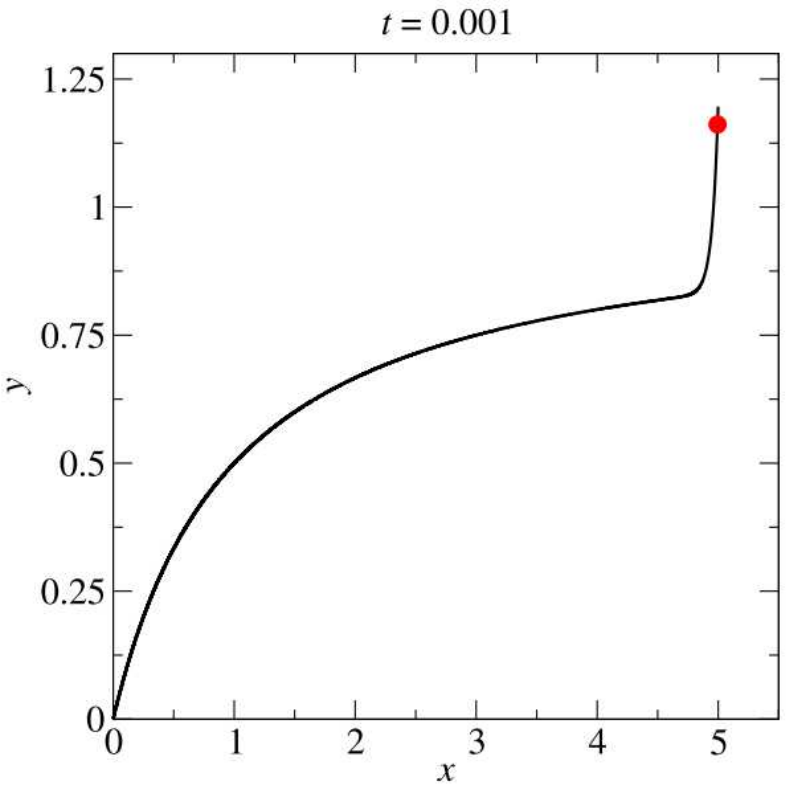


Time evolution of the state vector in the configuration space

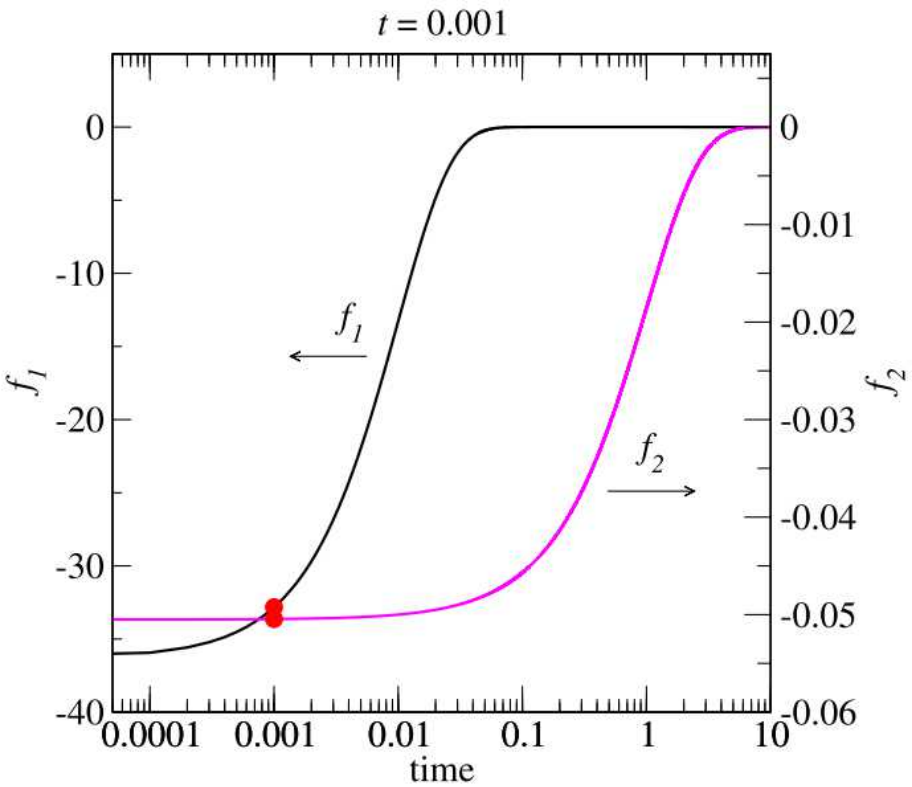


Time evolution of CSP (signed) mode amplitudes $f^i, i = 1, 2$

CSP Illustration with the Davis-Skodje ODE system

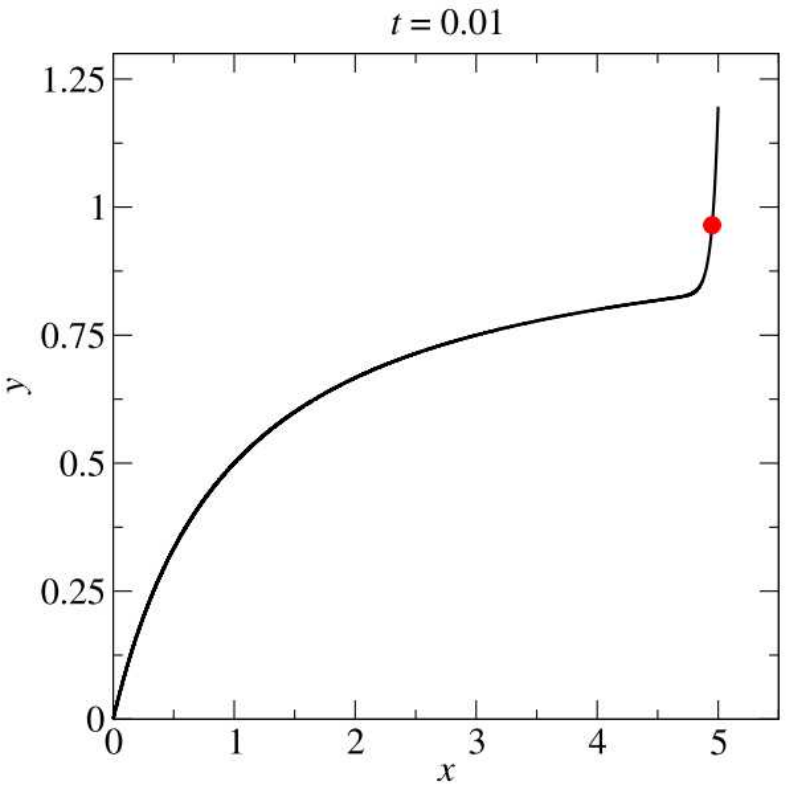


Time evolution of the state vector in the configuration space

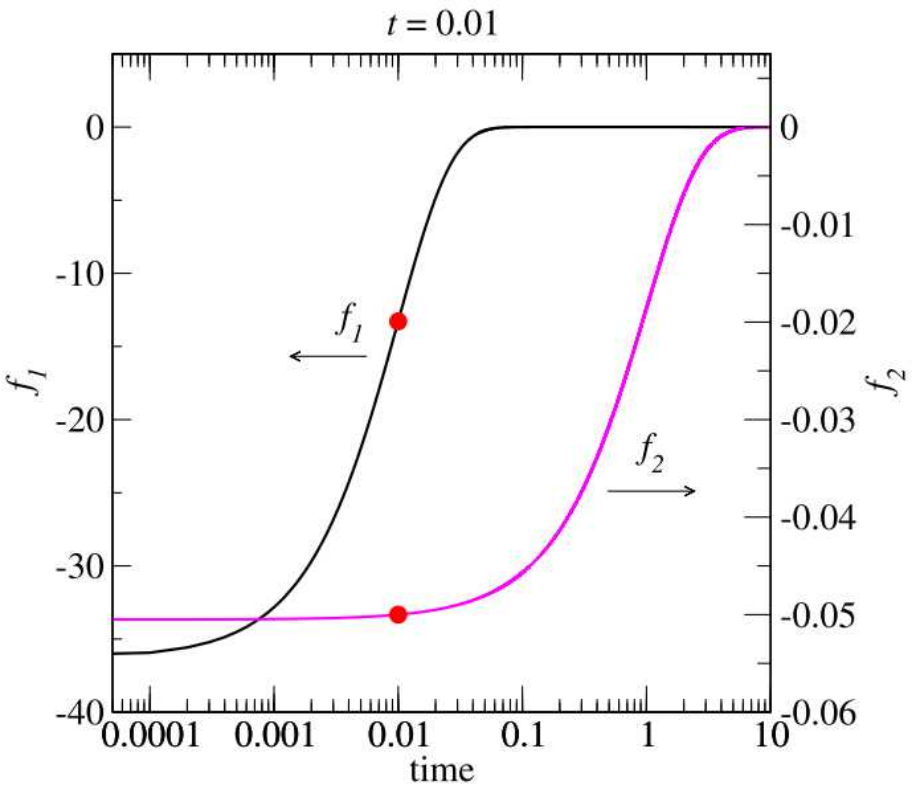


Time evolution of CSP (signed) mode amplitudes $f^i, i = 1, 2$

CSP Illustration with the Davis-Skodje ODE system

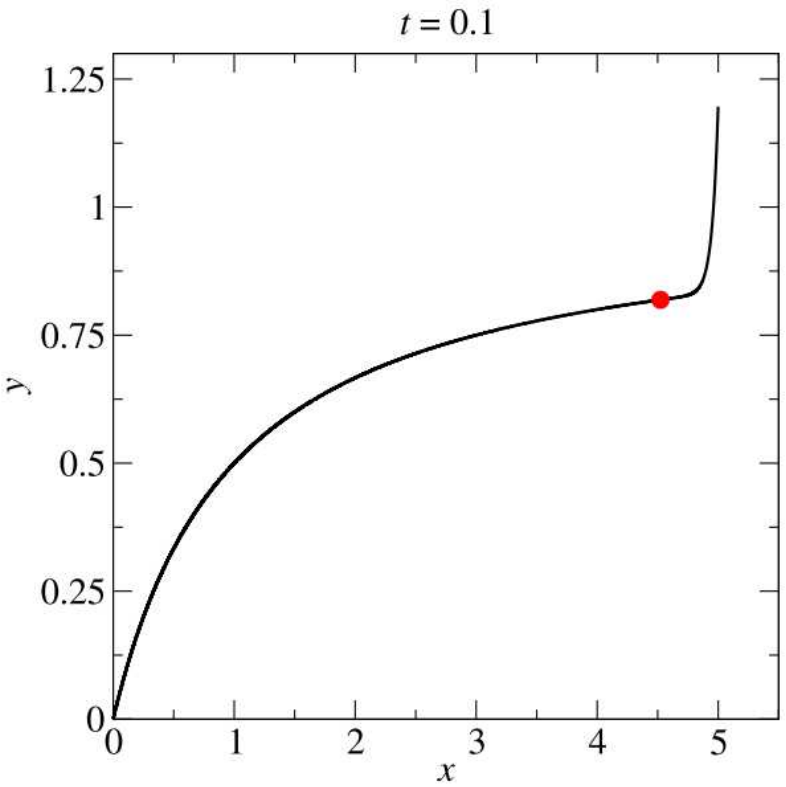


Time evolution of the state vector in the configuration space

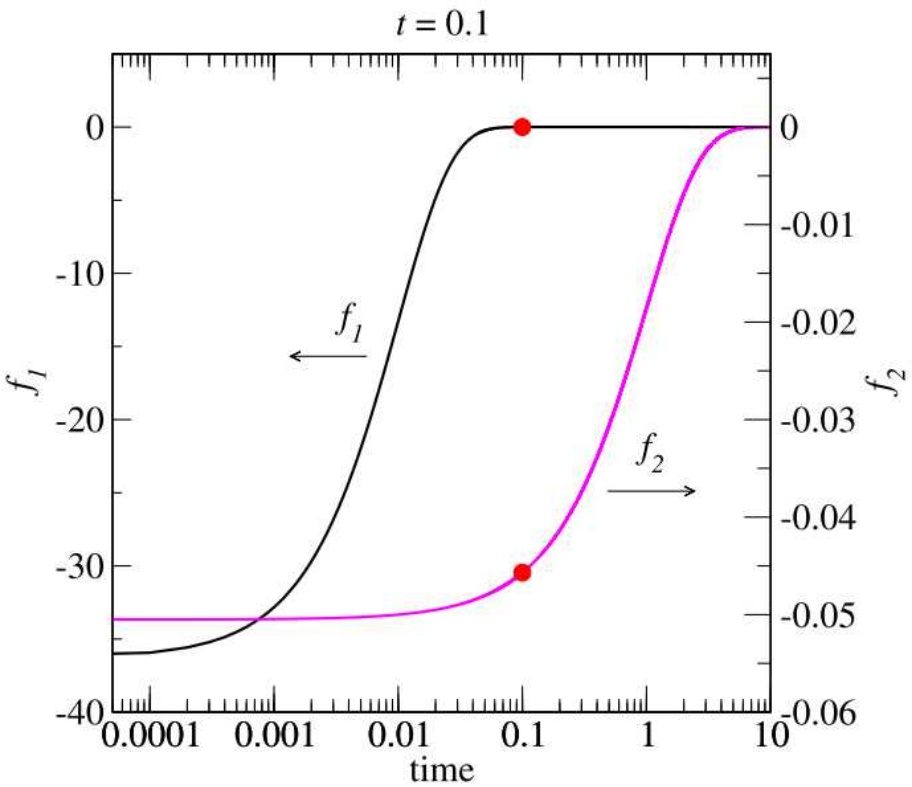


Time evolution of CSP (signed) mode amplitudes $f^i, i = 1, 2$

CSP Illustration with the Davis-Skodje ODE system

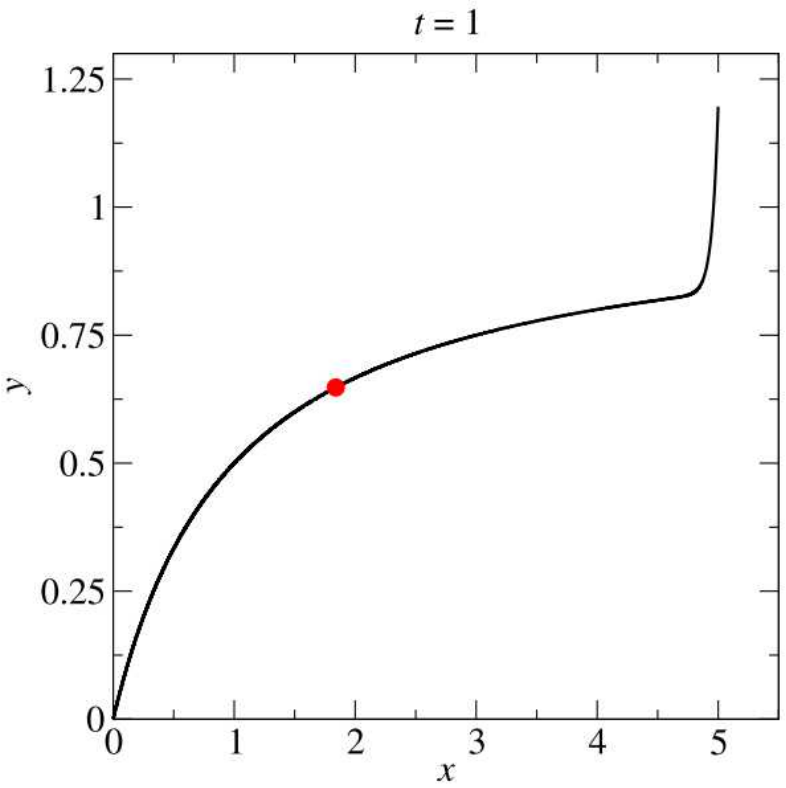


Time evolution of the state vector in the configuration space

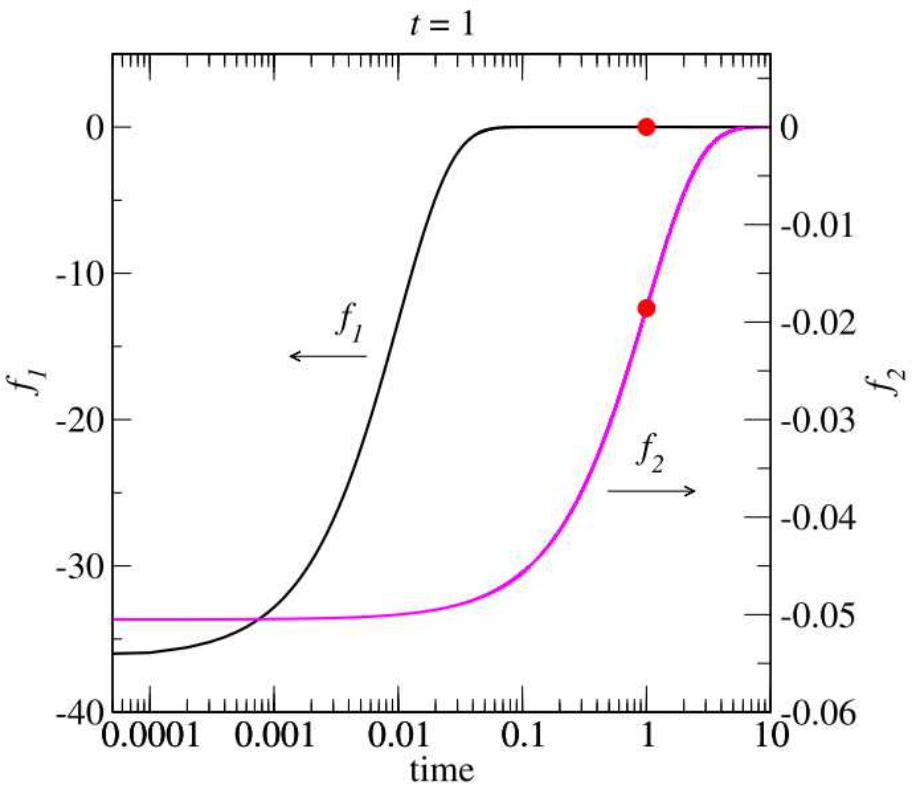


Time evolution of CSP (signed) mode amplitudes $f^i, i = 1, 2$

CSP Illustration with the Davis-Skodje ODE system

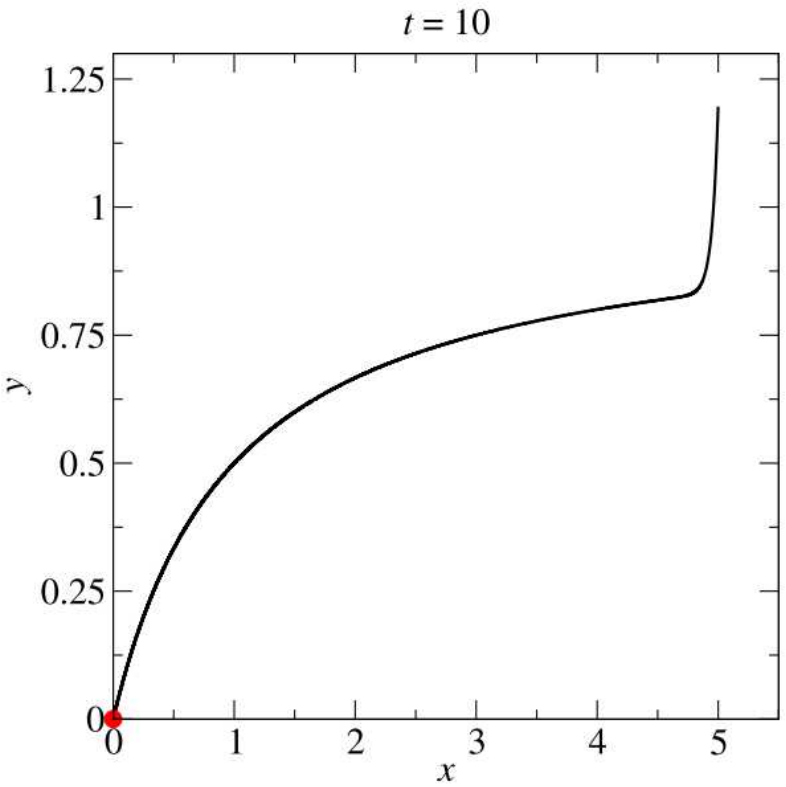


Time evolution of the state vector in the configuration space

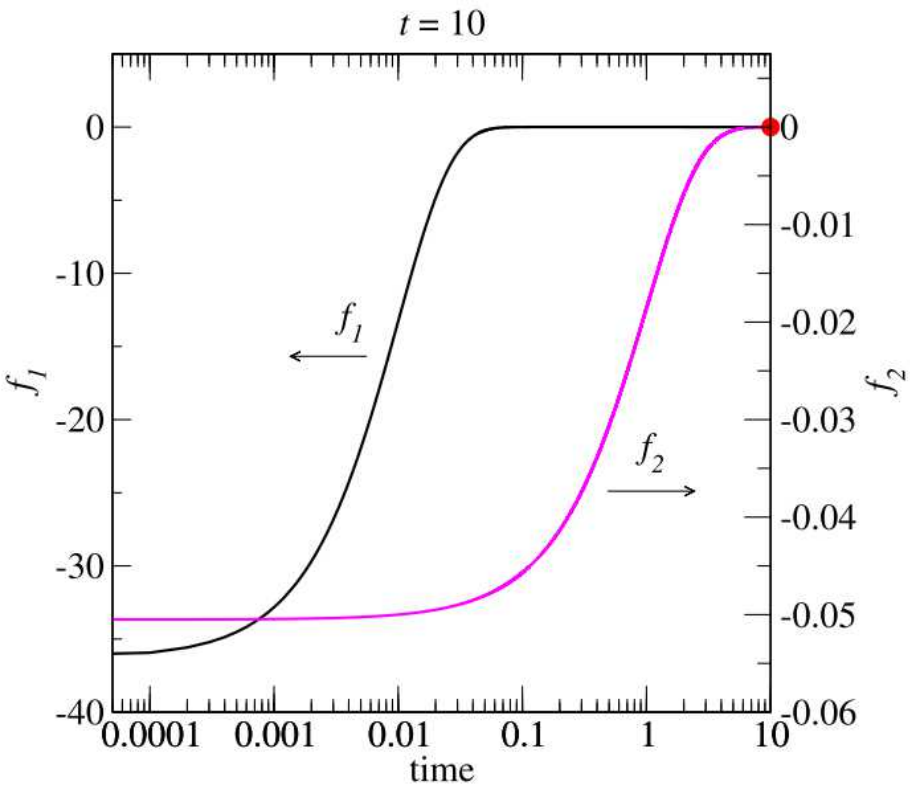


Time evolution of CSP (signed) mode amplitudes $f^i, i = 1, 2$

CSP Illustration with the Davis-Skodje ODE system



Time evolution of the state vector in the configuration space



Time evolution of CSP (signed) mode amplitudes $f^i, i = 1, 2$

CSP Decomposition – Fast (Exhausted) Subspace

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{g}(\boldsymbol{x}), \qquad \boldsymbol{x}(t=0) = \boldsymbol{x}_0$$

Evaluate the eigensolution for the Jacobian matrix $J_{\boldsymbol{g}} = \partial \boldsymbol{g}/\partial \boldsymbol{x}$, and sort the eigenmodes with negative real eigenvalue components in order of decreasing amplitude,

$$\lambda_1, \lambda_2, \cdots, \lambda_N \qquad \text{with } |\lambda_i| \geq |\lambda_{i+1}|$$

with time scales $\tau_i = 1/|\lambda_i|$ and $\tau_i \leq \tau_{i+1}$

$$\boldsymbol{g} = \sum_{i=1}^{N} \boldsymbol{a}_i f^i = \underbrace{\boldsymbol{a}_1 f^1 + \cdots + \boldsymbol{a}_M f^M}_{\boldsymbol{g}_{\text{fast}} \approx 0} + \underbrace{\boldsymbol{a}_{M+1} f^{M+1} + \cdots + \boldsymbol{a}_N f^N}_{\boldsymbol{g}_{\text{slow}}}$$

$$\boldsymbol{g}_{\text{slow}} = \sum_{s=M+1}^{N} \boldsymbol{a}_s f^s = \left(I - \sum_{r=1}^{M} \boldsymbol{a}_r \boldsymbol{b}^r \right) \boldsymbol{g} = P\boldsymbol{g}$$

CSP ODE Integrator

- Explicit integration in time, with $\Delta t = t_{n+1} - t_n = \mathcal{O}(\tau_{M+1})$

$$\boldsymbol{x}^{n+1} = \boldsymbol{x}^n + \sum_{i=1}^{M} \int_{t_n}^{t_{n+1}} \boldsymbol{a}_i(\boldsymbol{x}(t)) f^i(\boldsymbol{x}(t)) \mathrm{d}t + \sum_{i=M+1}^{N} \boldsymbol{a}_i(\boldsymbol{x}^n) f^i(\boldsymbol{x}^n) \Delta t$$

- Amplitudes of fast exhausted modes decay exponentially

$$f^i(t) \approx f^i(\boldsymbol{x}^n) \exp(-(t - t_n)/\tau_i(\boldsymbol{x}^n)), \qquad i = 1, \ldots, M$$

and, ignoring time variation of the basis vectors within Δt, we have

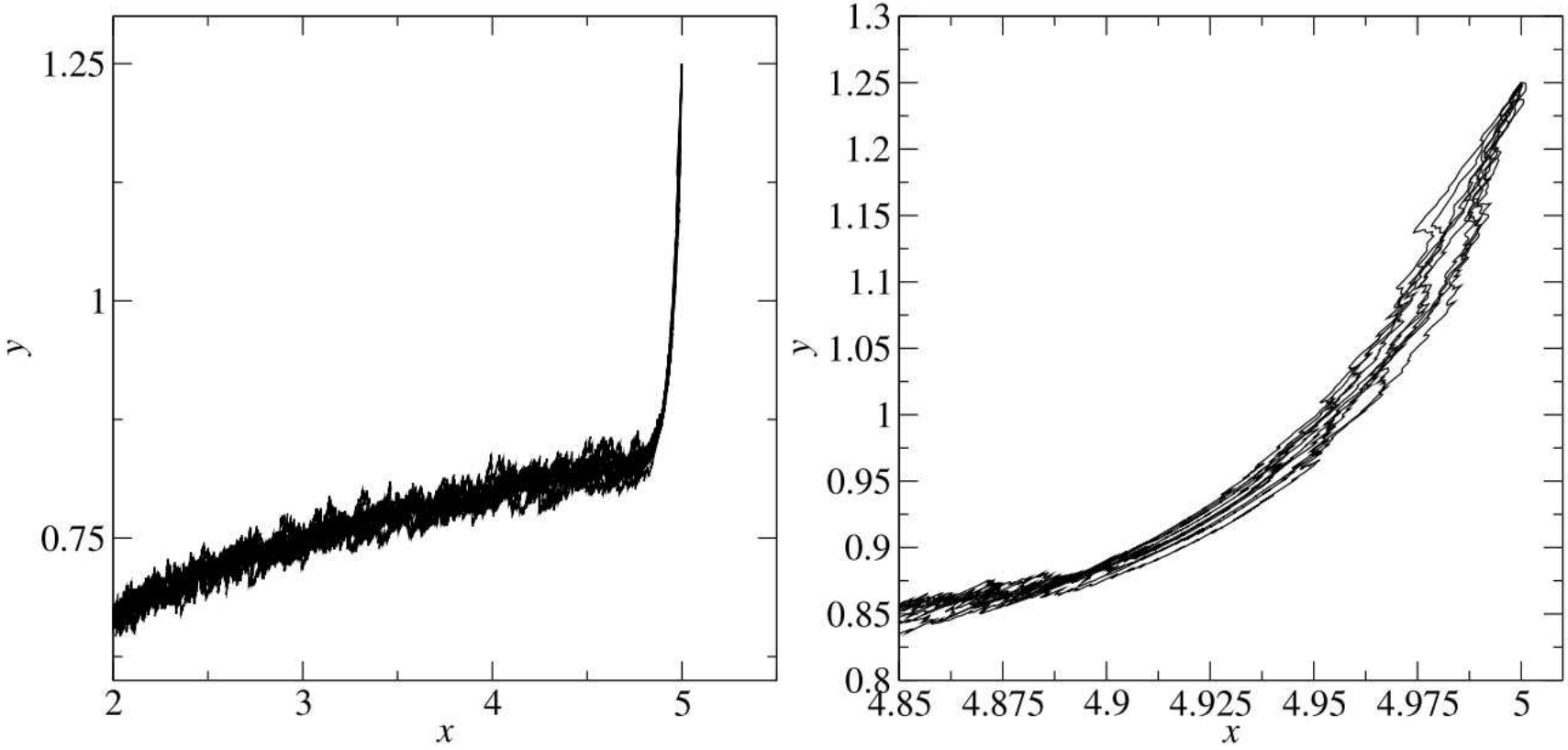
$$\begin{aligned} \int_{t_n}^{t_{n+1}} \boldsymbol{a}_i(\boldsymbol{x}(t)) f^i(\boldsymbol{x}(t)) \mathrm{d}t &\approx \boldsymbol{a}_i(\boldsymbol{x}^n) f^i(\boldsymbol{x}^n) \int_{t_n}^{t_{n+1}} e^{-(t-t_n)/\tau_i(\boldsymbol{x}^n)} \mathrm{d}t \\ &= \boldsymbol{a}_i(\boldsymbol{x}^n) f^i(\boldsymbol{x}^n) \tau_i(\boldsymbol{x}^n)(1 - e^{-\Delta t/\tau_i(\boldsymbol{x}^n)}) \end{aligned}$$

thus:

$$\boldsymbol{x}^{n+1} = \boldsymbol{x}^n + \sum_{i=1}^{M} \boldsymbol{a}_i^n f_n^i \tau_i^n (1 - e^{-\Delta t/\tau_i^n}) + \sum_{i=M+1}^{N} \boldsymbol{a}_i^n f_n^i \Delta t$$

What is the dynamical character of a stiff SDE?

- The drift term induces fast motion *in the mean* towards a stochastic manifold
- The diffusion term induces random motions, with zero mean, and no preferrential direction



CSP Applied to the CLE – 1

- CLE:

$$\boldsymbol{X}_{t+\mathrm{d}t} = \boldsymbol{X}_t + \boldsymbol{f}(\boldsymbol{X}_t)\mathrm{d}t + \sum_{j=1}^{R} \boldsymbol{g}_j(\boldsymbol{X}_t)\mathrm{d}W_j(t),$$

- Introduce the CSP basis $\{\boldsymbol{\alpha}_i, \boldsymbol{\beta}^i\}$, $i = 1, \ldots, N$, with $\boldsymbol{\beta}^i \cdot \boldsymbol{\alpha}_i = \delta_{ij}$, $\forall i$
- The (signed) mode amplitudes for the drift term are

$$\xi^i(\boldsymbol{X}) = \boldsymbol{\beta}^i(\boldsymbol{X}) \cdot \boldsymbol{f}(\boldsymbol{X})$$

with

$$\mathrm{d}\xi^i(t) = \mathrm{d}\boldsymbol{\beta}^i(t) \cdot \boldsymbol{f}(\boldsymbol{X}_t) + \boldsymbol{\beta}^i(t) \cdot \mathrm{d}\boldsymbol{f}(t)$$

CSP Applied to the CLE – 2

Using the stochastic chain rule, with some algebra, we have

$$\mathrm{d}\boldsymbol{f}(t) = \boldsymbol{\mu}(\boldsymbol{X}_t)\mathrm{d}t + \sum_{j=1}^{R} \boldsymbol{\sigma}_j(\boldsymbol{X}_t)\mathrm{d}W_j(t),$$

where

$$\mu_i(\boldsymbol{X}_t) = \sum_{k=1}^{N}\sum_{j,m=1}^{R} \nu_{jk}\nu_{mi}\rho_j(\boldsymbol{X}_t)\Big(\frac{\partial \rho_m(\boldsymbol{X}_t)}{\partial X_k} + \frac{1}{2}\sum_{l=1}^{N}\nu_{jl}\frac{\partial^2 \rho_m(\boldsymbol{X}_t)}{\partial X_k \partial X_l}\Big)$$

$$\sigma_{ij}(\boldsymbol{X}_t) = \sqrt{\rho_j(\boldsymbol{X}_t)}\sum_{k=1}^{N}\nu_{jk}\sum_{m=1}^{R}\nu_{mi}\frac{\partial \rho_m(\boldsymbol{X}_t)}{\partial X_k}$$

The change of ξ^i can be written as

$$\mathrm{d}\xi^i(t) = \left(\frac{\mathrm{d}\beta^i(t)}{\mathrm{d}t}\sum_{k=1}^{N}\xi^k\boldsymbol{\alpha}_k + \beta^i(t)\boldsymbol{\mu}(\boldsymbol{X}_t)\right)\mathrm{d}t + \beta^i(t)\sum_{j=1}^{R}\boldsymbol{\sigma}_j(\boldsymbol{X}_t)\mathrm{d}W_j(t)$$

CSP Applied to the CLE – 3

which reduces to:

$$\mathrm{d}\begin{pmatrix} \xi^1 \\ \vdots \\ \xi^N \end{pmatrix} = \Lambda(\boldsymbol{X}_t)\begin{pmatrix} \xi^1 \\ \vdots \\ \xi^N \end{pmatrix}\mathrm{d}t + \boldsymbol{\varphi}(\boldsymbol{X}_t)\mathrm{d}t + \Gamma(\boldsymbol{X}_t)\mathrm{d}\boldsymbol{W}(t),$$

$\boldsymbol{W}(t) = (W_1(t), \ldots, W_R(t))^T \in \mathbb{R}^{R\times 1}$ is an R-dimensional Brownian motion, and

$$\Lambda = \begin{bmatrix} (\frac{\mathrm{d}\boldsymbol{\beta}^1}{\mathrm{d}t} + \boldsymbol{\beta}^1 J_{\boldsymbol{f}})\boldsymbol{\alpha}_1 & \cdots & \cdots & (\frac{\mathrm{d}\boldsymbol{\beta}^1}{\mathrm{d}t} + \boldsymbol{\beta}^1 J_{\boldsymbol{f}})\boldsymbol{\alpha}_N \\ \vdots & \cdots & \cdots & \vdots \\ (\frac{\mathrm{d}\boldsymbol{\beta}^N}{\mathrm{d}t} + \boldsymbol{\beta}^N J_{\boldsymbol{f}})\boldsymbol{\alpha}_1 & \cdots & \cdots & (\frac{\mathrm{d}\boldsymbol{\beta}^N}{\mathrm{d}t} + \boldsymbol{\beta}^N J_{\boldsymbol{f}})\boldsymbol{\alpha}_N \end{bmatrix} \in \mathbb{R}^{N\times N}$$

$$\boldsymbol{\varphi} = \begin{bmatrix} \boldsymbol{\beta}^1 \\ \vdots \\ \boldsymbol{\beta}^N \end{bmatrix} \cdot \begin{bmatrix} \sum_{k,l=1}^{N}\sum_{j,m=1}^{R} \nu_{jk}\nu_{jl}\nu_{m1}\frac{\partial^2\rho_m}{\partial X_k \partial X_l}\rho_j \\ \vdots \\ \sum_{k,l=1}^{N}\sum_{j,m=1}^{R} \nu_{jk}\nu_{jl}\nu_{mN}\frac{\partial^2\rho_m}{\partial X_k \partial X_l}\rho_j \end{bmatrix} \in \mathbb{R}^{N\times 1}$$

$$\Gamma = \begin{bmatrix} \boldsymbol{\beta}^1 \\ \vdots \\ \boldsymbol{\beta}^N \end{bmatrix} \cdot \begin{bmatrix} \sigma_{11} & \cdots & \cdots & \sigma_{1R} \\ \vdots & \cdots & \cdots & \vdots \\ \sigma_{N1} & \cdots & \cdots & \sigma_{NR} \end{bmatrix} \in \mathbb{R}^{N\times R}$$

CSP Applied to the CLE – 4

- Diagonalization of Λ is not sufficient to decouple fast and slow processes, due to $\boldsymbol{\varphi}$ and $\Gamma\,\mathrm{d}\boldsymbol{W}(t)$
- A refinement procedure to separate the fast and slow dynamics is necessary
- We use the above SDE, linearized, to motivate the use of the eigenvectors of J_f as CSP basis vectors, as done for the ODE case
- A linearized analysis gives

$$\mathrm{d}\begin{pmatrix} \xi^1 \\ \vdots \\ \xi^N \end{pmatrix} = \Lambda \begin{pmatrix} \xi^1 \\ \vdots \\ \xi^N \end{pmatrix} \mathrm{d}t + \Gamma(\boldsymbol{X}(t))\mathrm{d}\boldsymbol{W}(t)$$

  in which the time evolution of the mean modes for eigenvalues with different real parts are all separated
- NB. for the nonlinear CLE, as opposed to a general SDE, it is expected that the magnitudes of components of $\boldsymbol{\varphi}$ are small relative to those of Λ, thus the use of the linearized approximation is viable

CSP Applied to the CLE – 5

Proposed strategy:

- Stiffness in the drift term can be dealt with similar to the ODE RHS
  - Evaluate eigendecomposition of the Jacobian of the drift term at each time step
  - Identify fast/slow subspaces, **determine** $M \Leftarrow$ **main challenge**
  - Integrate fast processes in time by modeling their exponential decay to the manifold
  - Integrate slow processes using EM
- Integrate diffusion using EM

CLE-CSP Time Integration

- Path-wise CLE Difference equations, with $\boldsymbol{X}_t^\omega := \boldsymbol{X}(t, \boldsymbol{\eta}^\omega)$

$$\boldsymbol{X}^\omega_{t+\mathrm{d}t} = \boldsymbol{X}^\omega_t + \boldsymbol{f}(\boldsymbol{X}^\omega_t)\mathrm{d}t + \boldsymbol{g}((\boldsymbol{X}^\omega_t, \boldsymbol{\eta}^\omega)\sqrt{\mathrm{d}t}.$$

where

$$\boldsymbol{g}(\boldsymbol{X}^\omega_t, \boldsymbol{\eta}^\omega) := \sum_{j=1}^{R} \boldsymbol{g}_j(\boldsymbol{X}(t, \boldsymbol{\eta}^\omega))\eta^\omega_j(t)$$

$$\boldsymbol{\eta}^\omega(t) := (\eta^\omega_1(t), \ldots, \eta^\omega_R(t))^T$$

and $\mathcal{N}_j(t, \omega) := \eta^\omega_j(t)$ is the continuous sample path for $\mathcal{N}_j, \omega \in \Omega$

Thus the CSP time integration is as follows,

$$\boldsymbol{X}^\omega_{t+\Delta t} = \boldsymbol{X}^\omega_t + \sum_{i=1}^{M} \xi^i(\boldsymbol{X}^\omega_t)\boldsymbol{\alpha}_i(\boldsymbol{X}^\omega_t)\tau^F_i(t)(1 - e^{-\Delta t/\tau^F_i(t)}) + \sum_{i=M+1}^{N} \xi^i\boldsymbol{\alpha}_i\Delta t + \boldsymbol{g}((\boldsymbol{X}^\omega_t, \boldsymbol{\eta}^\omega)\sqrt{\Delta t}$$

CLE-CSP – Determination of M

- The determination of M in an ODE setting is as follows:

$$M = \text{max } m \quad \text{s.t.} \quad \tau_{m+1} \left| \sum_{i=1}^{m} \boldsymbol{a}_i f^i \right| < \epsilon_r \boldsymbol{x} + \epsilon_a \mathbf{1}$$

- Stochastic noise renders this test ineffectual for the CLE
- The mode amplitudes ξ^i for any sample-path do not decay to zero
- Choosing an arbitrary threshold is unreliable
- A reliable approach involves utilization of sample-path statistics

CLE-CSP – Determination of M – example

Consider the "VGCN" system involving the three species $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$:

(Valorani, 2005)

$$\mathcal{X} \underset{k_1^b}{\overset{k_1^f}{\rightleftharpoons}} 2\mathcal{Y}, \qquad \mathcal{X}+\mathcal{Y} \underset{k_2^b}{\overset{k_2^f}{\rightleftharpoons}} \mathcal{Z}, \qquad \mathcal{Y}+\mathcal{Z} \underset{k_3^b}{\overset{k_3^f}{\rightleftharpoons}} \mathcal{X}.$$

With the state vector $\boldsymbol{U}_t := (X_t, Y_t, Z_t)^T$, the path-wise CLEs:

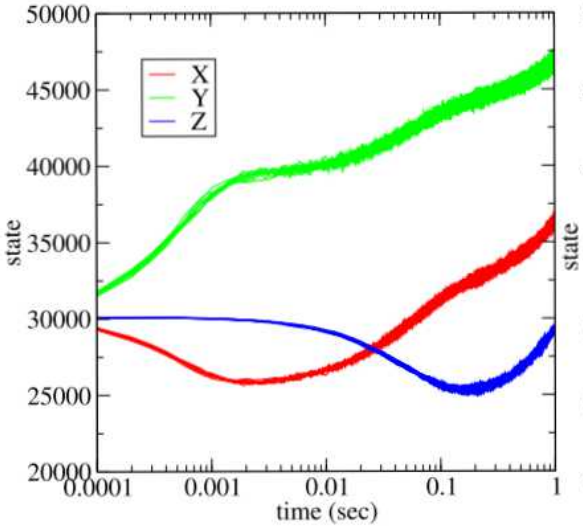
$$\mathrm{d}\boldsymbol{U}_t^\omega = \boldsymbol{f}(\boldsymbol{U}_t^\omega)\Delta t + \boldsymbol{g}(\boldsymbol{U}_t^\omega, \boldsymbol{\eta}^\omega)\sqrt{\Delta t},$$

where, employing the reaction propensities $(r_1^f, r_1^b, r_2^f, r_2^b, r_3^f, r_3^b)$,

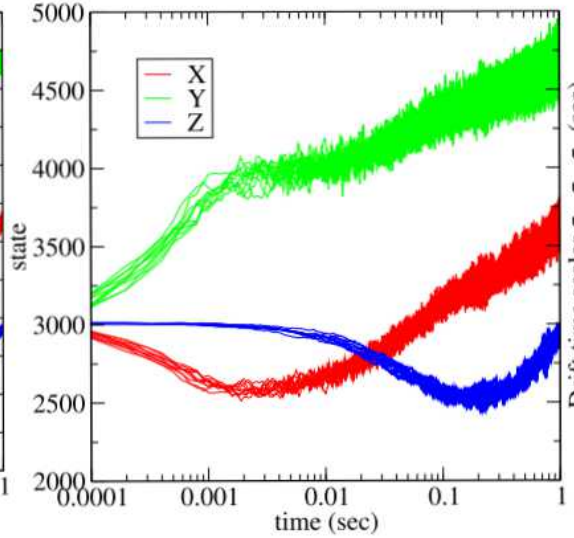
$$\boldsymbol{f} = \begin{pmatrix} -r_1^f X + r_1^b Y(Y-1) - r_2^f XY + r_2^b Z + r_3^f YZ - r_3^b X \\ r_1^f X - r_1^b Y(Y-1) - r_2^f XY + r_2^b Z - r_3^f YZ + r_3^b X \\ r_2^f XY - r_2^b Z - r_3^f YZ + r_3^b X \end{pmatrix}$$

$$\boldsymbol{g} = \begin{bmatrix} -\sqrt{r_1^f X} & \sqrt{r_1^b Y(Y-1)} & -\sqrt{r_2^f XY} & \sqrt{r_2^b Z} & \sqrt{r_3^f YZ} & -\sqrt{r_3^b X} \\ \sqrt{r_1^f X} & -\sqrt{r_1^b Y(Y-1)} & -\sqrt{r_2^f XY} & \sqrt{r_2^b Z} & -\sqrt{r_3^f YZ} & \sqrt{r_3^b X} \\ 0 & 0 & \sqrt{r_2^f XY} & -\sqrt{r_2^b Z} & -\sqrt{r_3^f YZ} & \sqrt{r_3^b X} \end{bmatrix} \begin{pmatrix} \eta_1^\omega(t) \\ \eta_2^\omega(t) \\ \eta_3^\omega(t) \\ \eta_4^\omega(t) \\ \eta_5^\omega(t) \\ \eta_6^\omega(t) \end{pmatrix}$$

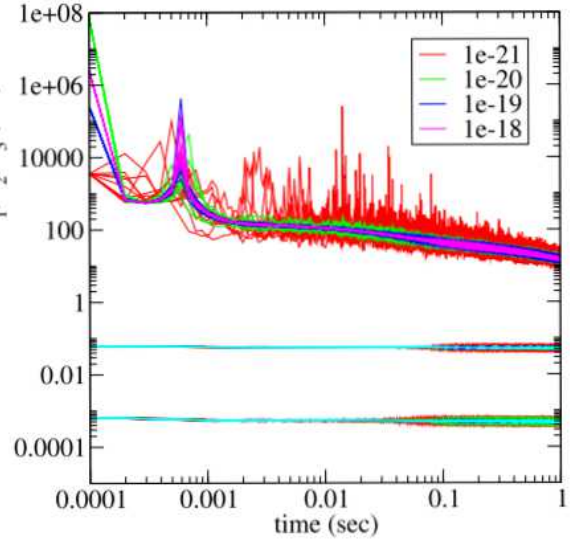
VGCN System Dynamics



vol=10^{-19} lt



vol=10^{-20} lt

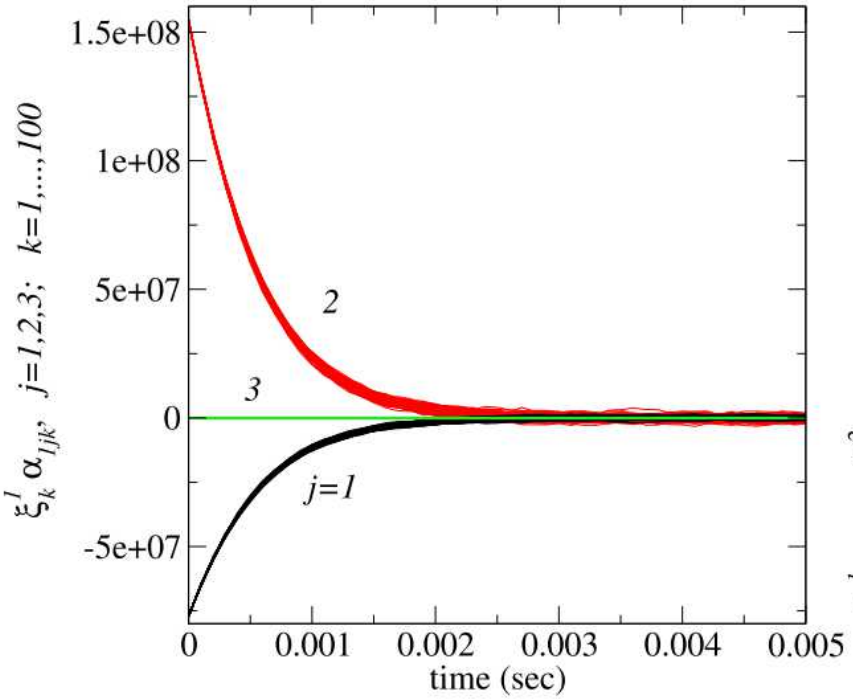


vol=$10^{-18} - 10^{-21}$ lt

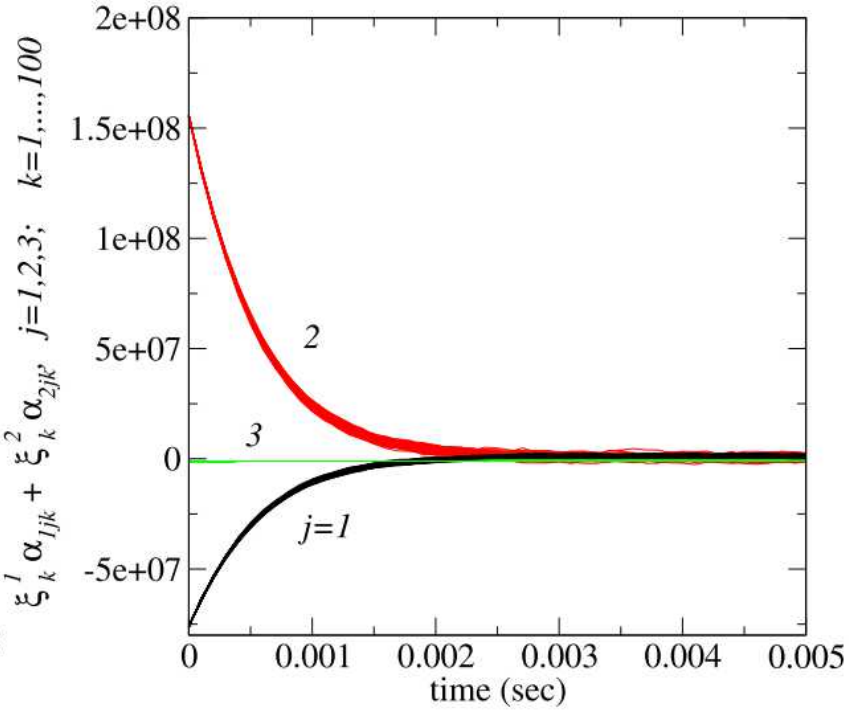
- System exhibits 3 drift-term negative real eigenvalues with associated time scales
- Mean time scales exhibit essentially no dependence on system size
- Decay to a 2D : 1D : 0D manifolds in succession

VGCN System Dynamics – Mode Contributions

100 samples



$\xi^1 \boldsymbol{\alpha}_1$

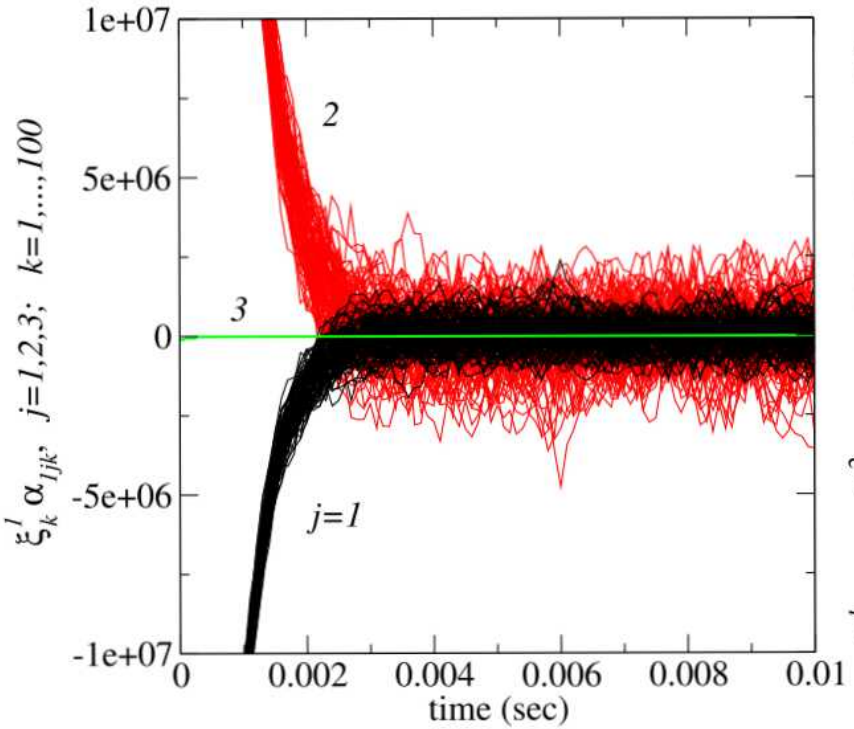


$\xi^1 \boldsymbol{\alpha}_1 + \xi^2 \boldsymbol{\alpha}_2$

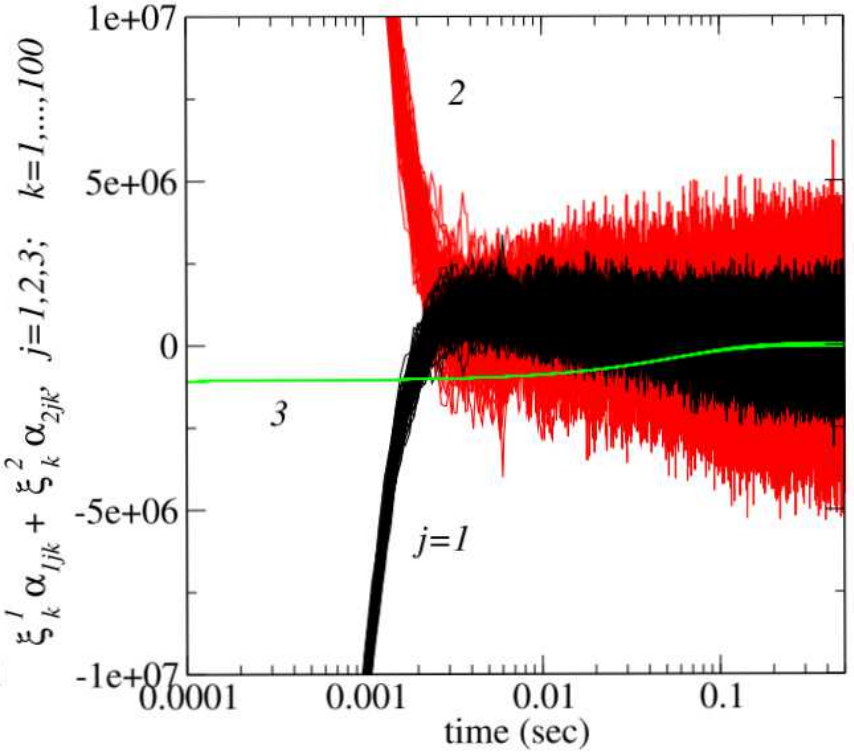
- Noise leads to challenging detection problem for when a set of modes is exhausted
- Need a robust means of selecting thresholds
- Ensure that the absolute value of the sums is small

VGCN System Dynamics – Mode Contributions

100 samples



$\xi^1 \boldsymbol{\alpha}_1$



$\xi^1 \boldsymbol{\alpha}_1 + \xi^2 \boldsymbol{\alpha}_2$

- Noise leads to challenging detection problem for when a set of modes is exhausted
- Need a robust means of selecting thresholds
- Ensure that the absolute value of the sums is small

A Reliable M-Detection Strategy

- Run K samples concurrently
- Examine statistics of the absolute value of the sum of mode contributions to the drift term

$$\mathfrak{S}_{jmk} = \Big| \sum_{i=1}^{m} \xi_k^i \alpha_{ijk} \Big|, \qquad j = 1, \ldots, N; \qquad k = 1, \ldots, K.$$

Define the N-long sample mean and standard deviation vectors μ_{mK} and σ_{mK}, where, for $j = 1, \ldots, N$

$$\begin{aligned} \mu_{jmK} &= \frac{1}{K} \sum_{k=1}^{K} \mathfrak{S}_{jmk} \\ \sigma_{jmK} &= \Big[\frac{1}{K-1} \sum_{k=1}^{K} (\mathfrak{S}_{jmk} - \mu_{jmK})^2 \Big]^{1/2} \end{aligned}$$

A Reliable M-Detection Strategy

- Declare a set of modes exhausted when their above sample-mean is small relative to the associated sample standard deviation
- Also ensure:
  - This set does not include a mode with a positive real eigenvalue component
  - The drift time scale of the fastest slow mode ("active" mode) is faster than the fastest diffusion time scale

Thus:

$$M = \max m \in [1, N]$$

s.t.

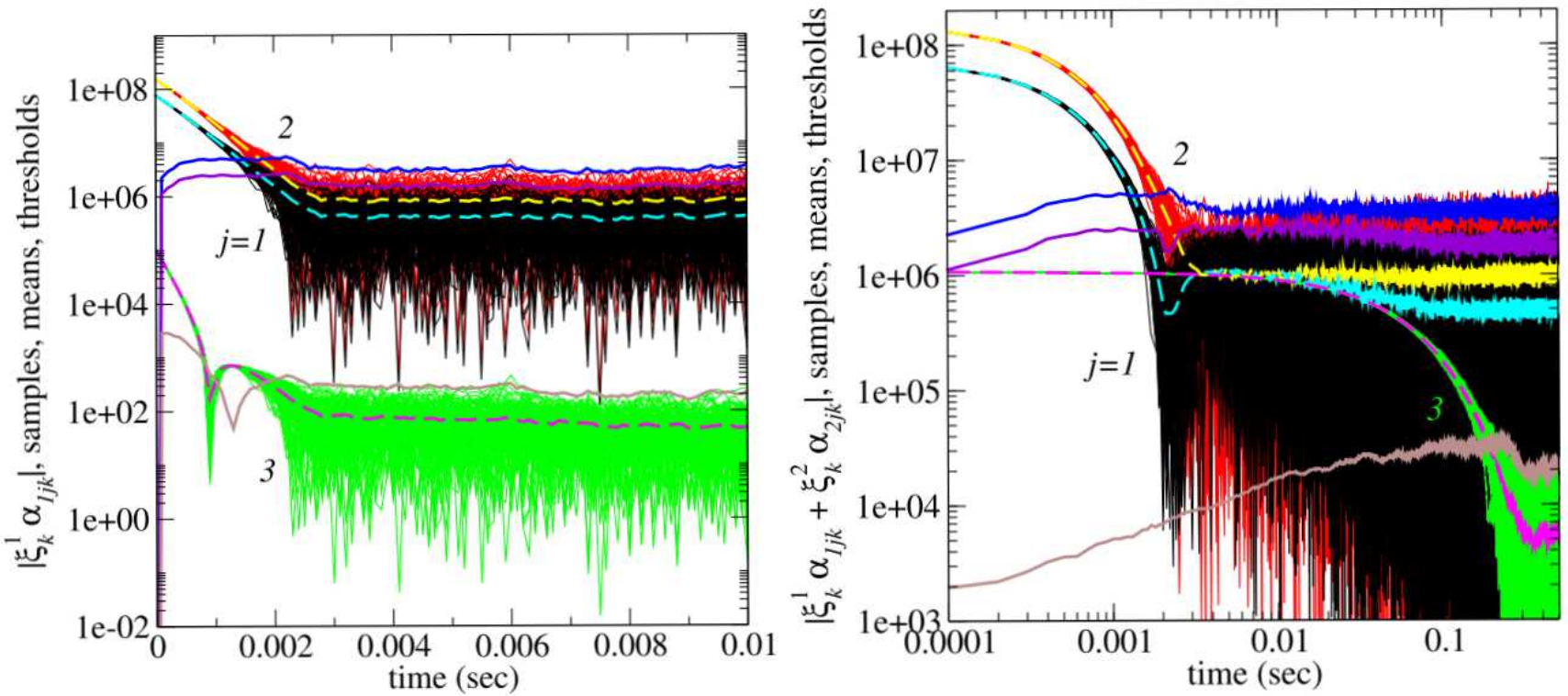
$$\mu^F_{jmK} < \beta\ \sigma^F_{jmK}, \quad \forall j \in [1, N]$$

$$\max_k \tau^F_{M+1,k} < \gamma\ \min_k \tau^G_{1,k}$$

$$\mathrm{Re}(\lambda^F_{rk}) < 0, \quad \forall r \in [1, m]; \quad \forall k \in [1, K]$$

We use $\beta = 5, \gamma = 0.5$.

M-Detection in the VGCN system



- Dashed lines are means, solid lines are γ-scaled standard deviations
- Mode 1 is exhausted at $\sim$1.6 msec $\quad \tau_1^F = 0.6$ msec
- Mode 2 is exhausted at $\sim$0.15 sec $\quad \tau_2^F = 0.06$ sec

Choice of time step

- Integrate all K samples synchronously in time, same Δt for all

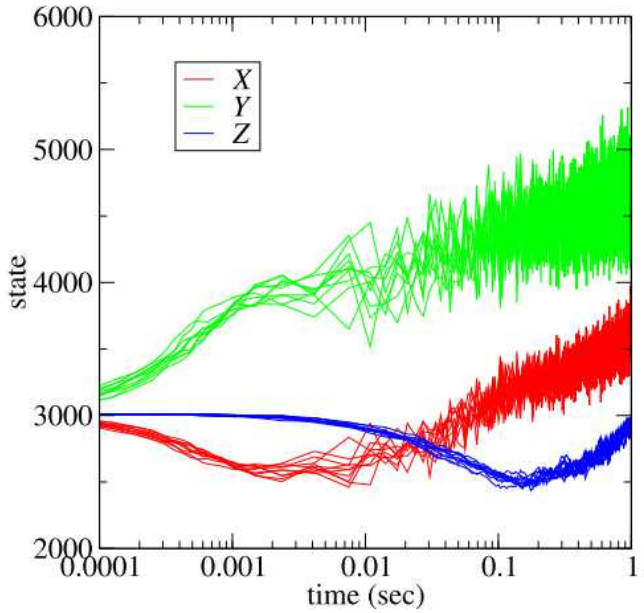
$$\Delta t^* = \begin{cases} 0.1 \min_k \tau^F_{1,k} & \text{for} \quad M = 0 \\ (1-\theta) \max_k \tau^F_{M,k} + \theta \min_k \tau^F_{M+1,k} & \text{for} \quad M > 0 \end{cases}$$

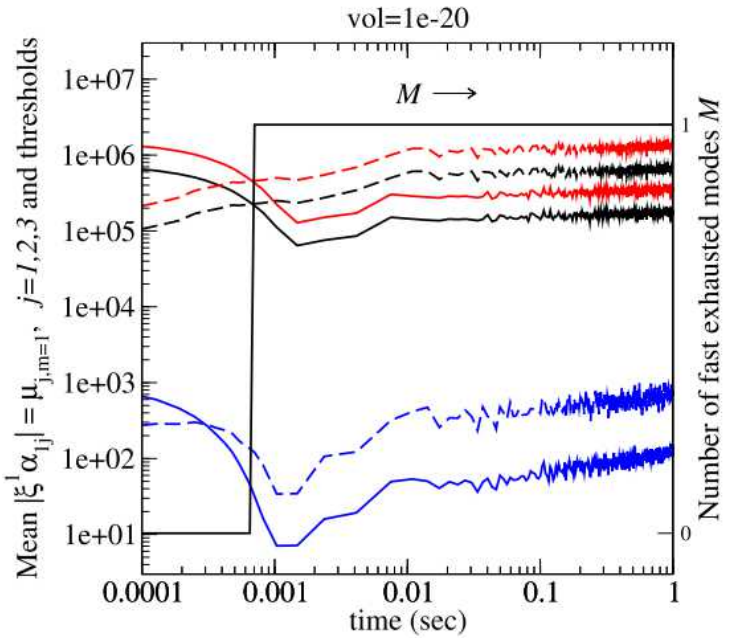
  where $0 < \theta < 1$, is a chosen constant. We use $\theta = 0.05$.

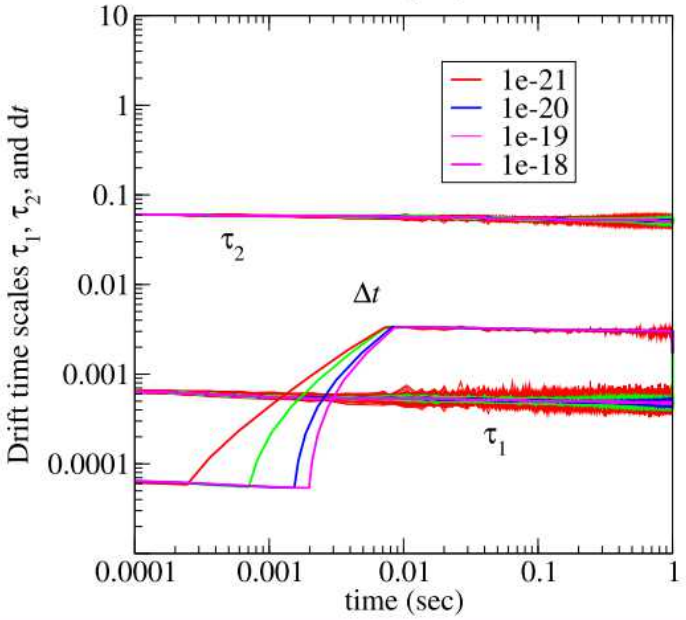
- Further, for minimizing the impact of large sudden increases in the time-step size, we enforce

$$\Delta t_{\mathrm{n}} = \min(\Delta t^*, 2\Delta t_{\mathrm{n}-1})$$

VGCN CLE time integration with CSP

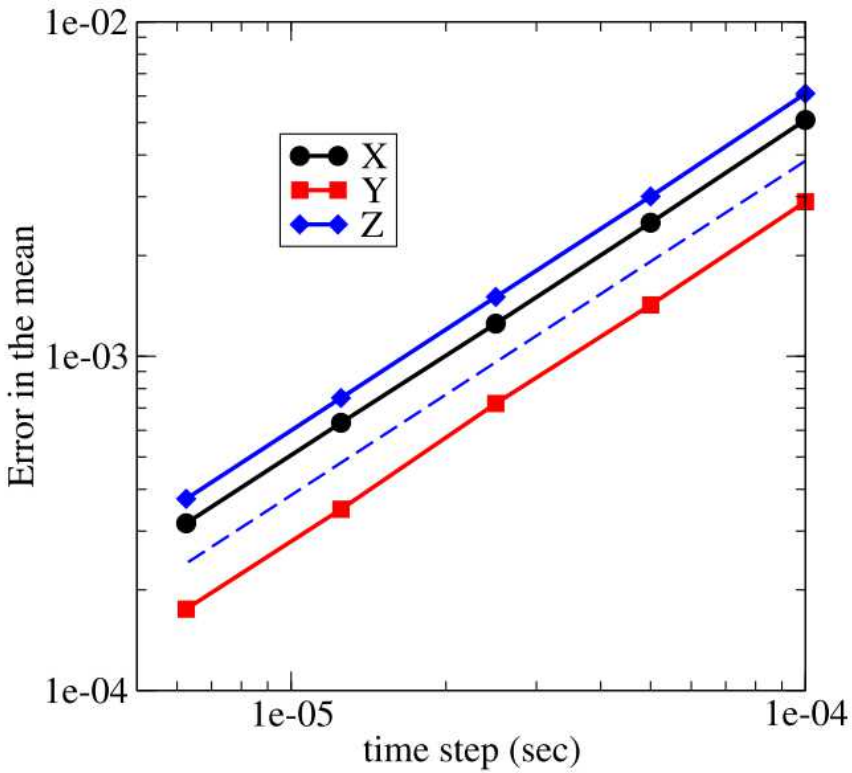


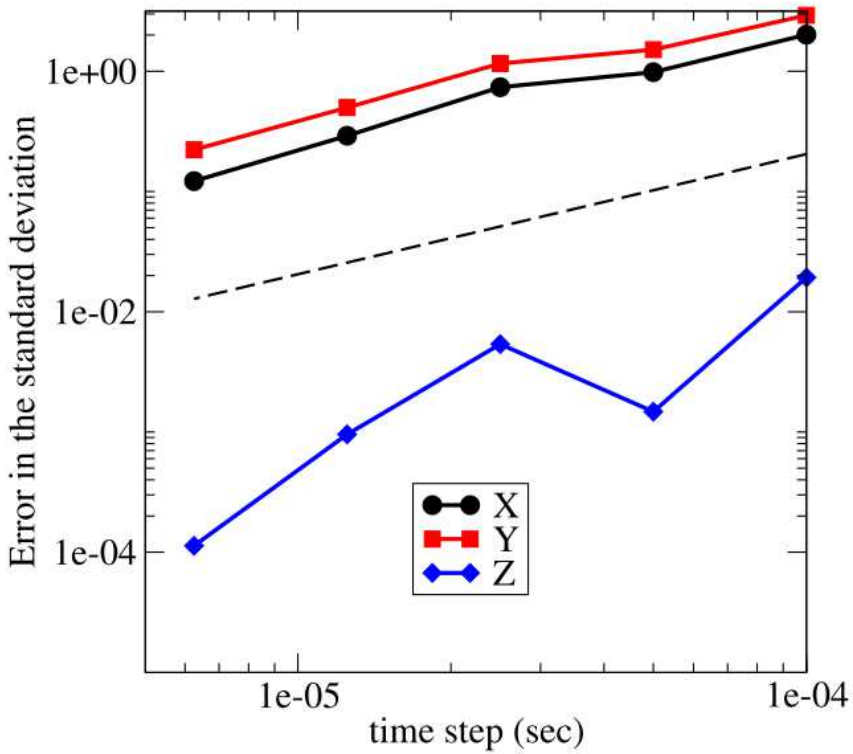




- Qualitatively similar evolution of the state variables
- M selection goes up to $M = 1$
  - Limited by the fastest diffusion timescale
- Time step increase by nearly two orders of magnitude

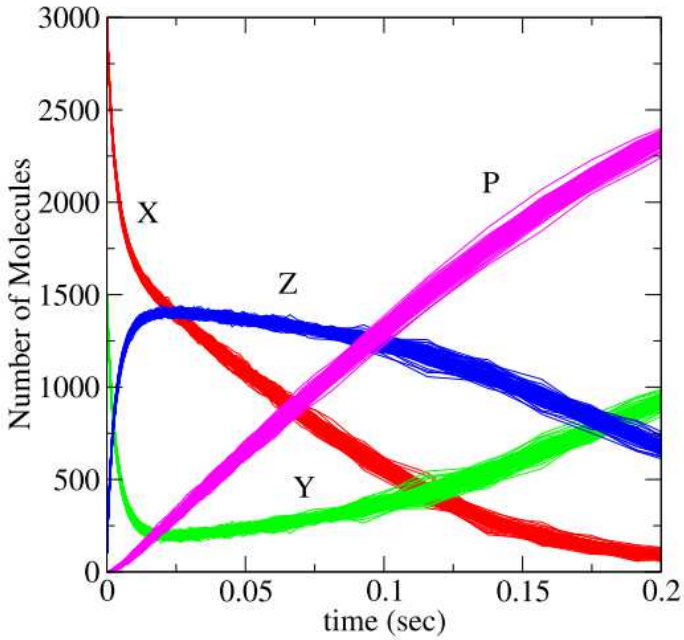
VGCN CLE time integration with CSP – Error convergence

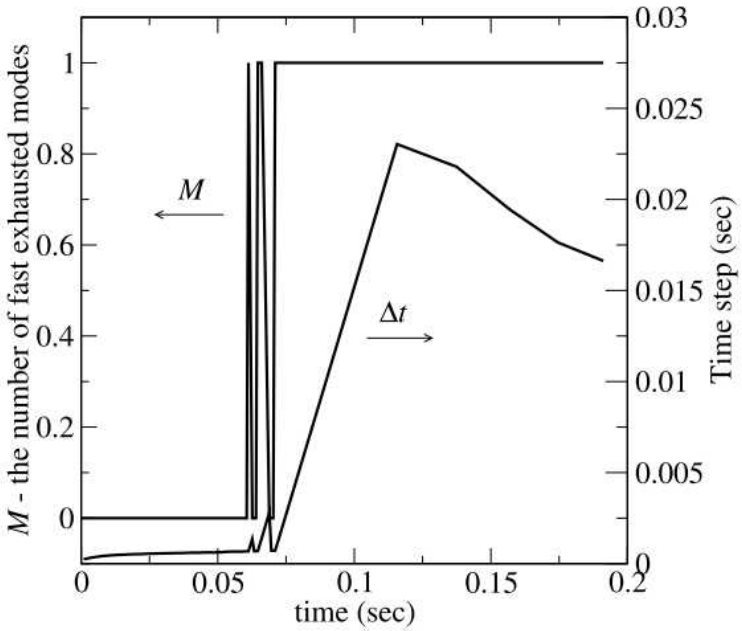




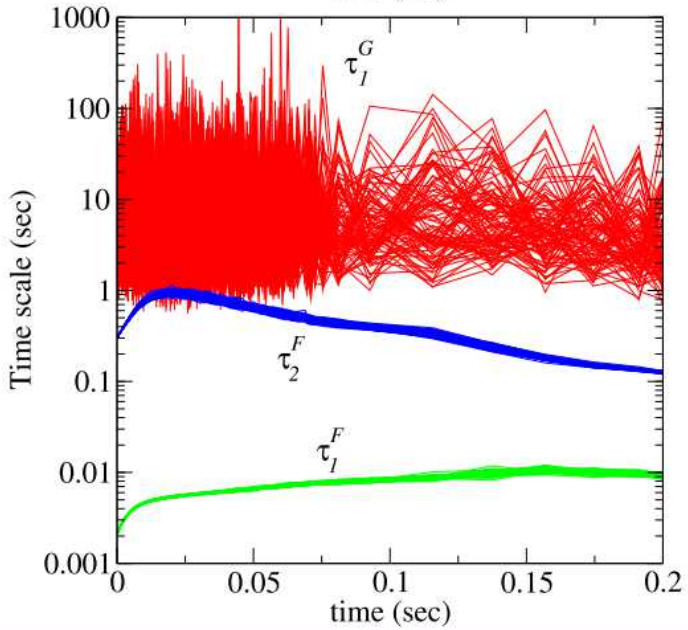
- First order weak-convergence of EM is retained, for both the mean and standard deviation
- Convergence is with respect to small-Δt computations with EM

Michaelis-Menten system CLE

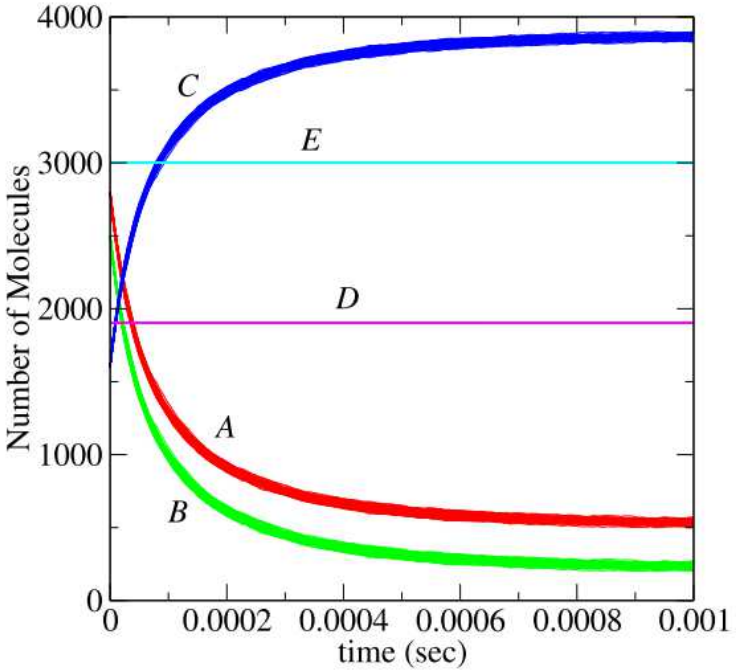


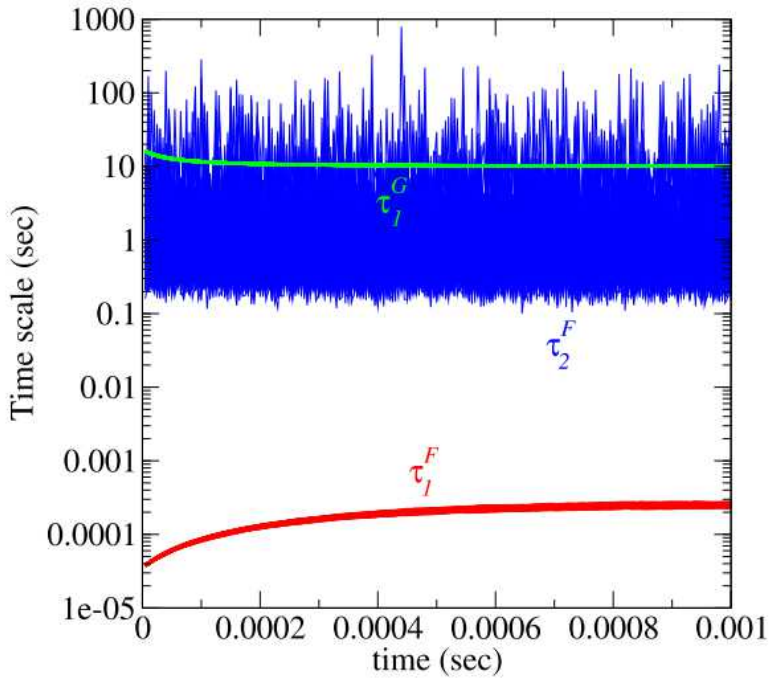


- M goes up to a maximum of 1
- Limited by diffusion time scales



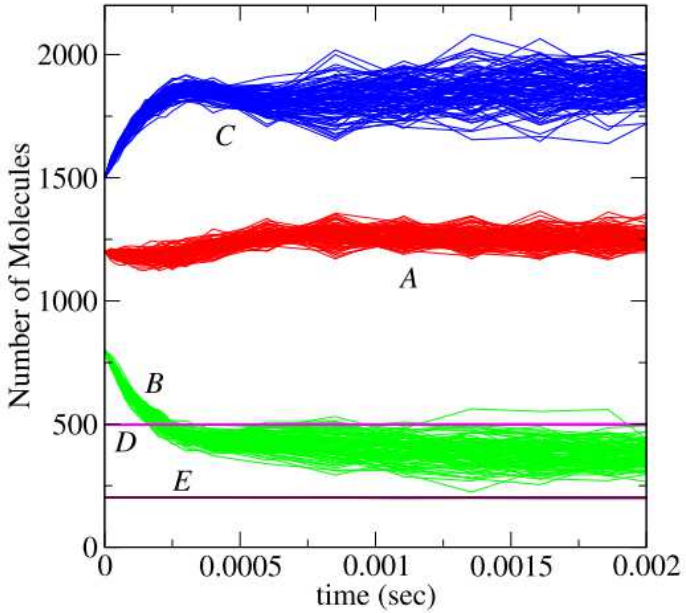
Protein system CLE

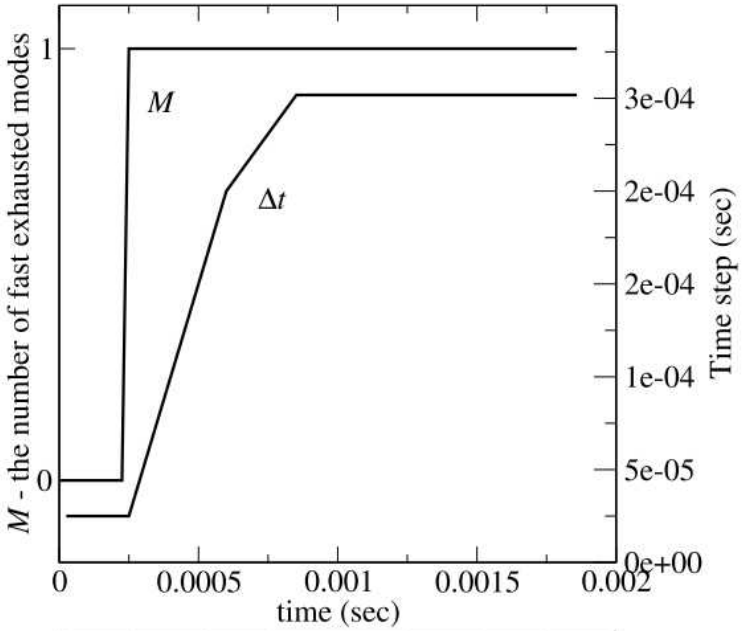




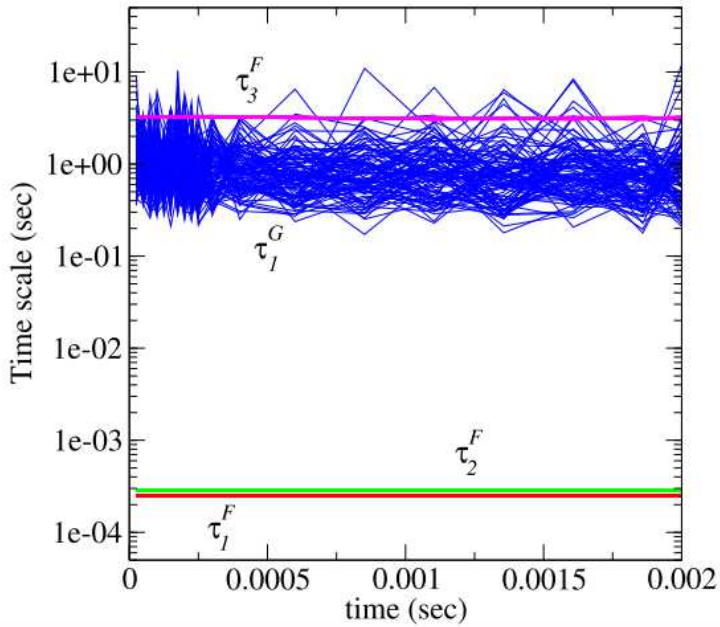
- Here $M = 0$ given the fast diffusional time scales

Acyclic system CLE





- M goes up to a maximum of 1
- Limited by diffusion time scales



Computational Performance

- The computational savings of large explicit time steps have to be balanced against the costs of Jacobian eigensolves.
  - Currently, the EM implementation is 1.5$\times$ faster than the CSP integration
- Potential remedies to improve computational performance include
  - Resolving the diffusion-induced upper limit on Δt, thus allowing larger time step computations
  - Reusing the computed eigendecomposition of the Jacobian over some number of time steps
  - Exploring eigensolvers that can
    - make efficient use of a good initial guess
    - compute only the fastest $M + 1$ eigenmodes

Closure

- We demonstrated the utility of CSP for enabling large time step explicit integration of stiff CLEs
- Numerous directions for future work are feasible
  - Develop adequate modeling of fast diffusional processes
  - Reduce eigendecomposition costs
    - reuse, good initial guess, partial eigensolve
  - Pursue theoretical convergence proofs
  - Development of treatment for multiple manifolds and switching between basins of attraction