Computing the quasipotential for nongradient SDEs

Maria Cameron

Joint work with

Daisy Dahiya,

Shuo Yang,

Samuel Potter,

Nicholas Paskal

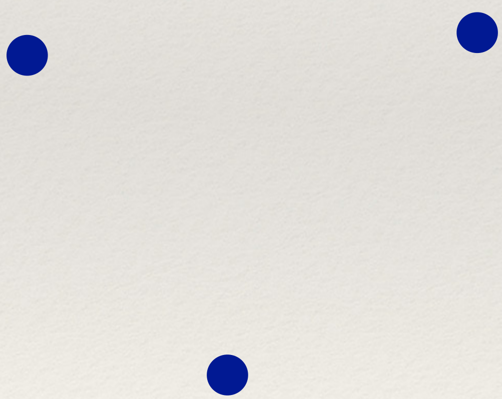
University of Maryland, College Park, MD, USA

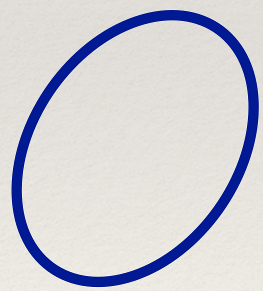
RESIM-2021

What are rare events?

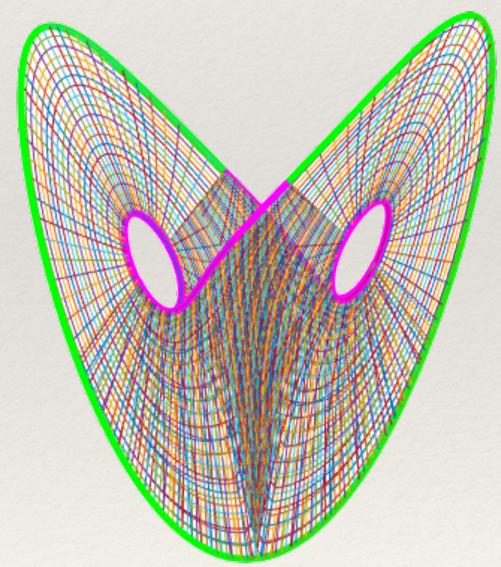
$$dX_t = \underbrace{b(X_t)}_{\text{C^1 deterministic vector field}}dt + \sigma(X_t)\underbrace{\sqrt{\epsilon}}_{\text{small parameter}}\underbrace{dw_t}_{\text{Brownian motion}}$$

Stable equilibria





Stable limit cycle



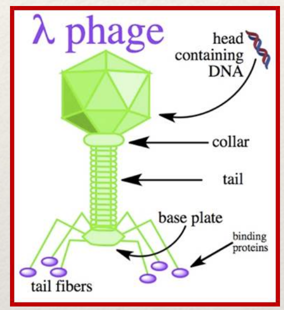
More complex attractors

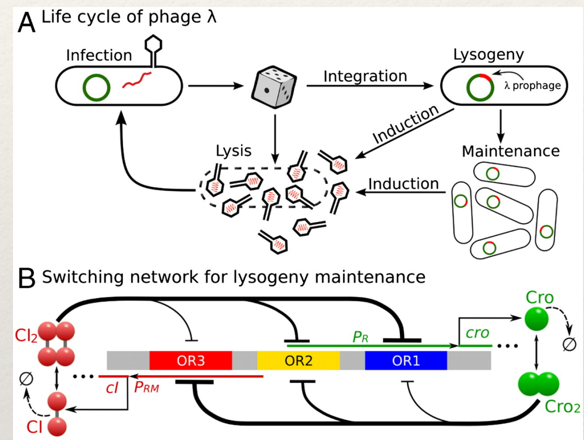
Where nongradient SDE models come from

$$dX_t = b(X_t)dt + \sigma(X_t)\sqrt{\epsilon}dw_t$$

Biological and ecological models

- Genetic switches
  - Lambda Phage (Shea et al. (1980s), Aurell and Sneppel (2002)), 2D
  - Two-state gene expression model with positive feedback (Lv et al. 2014), 3D

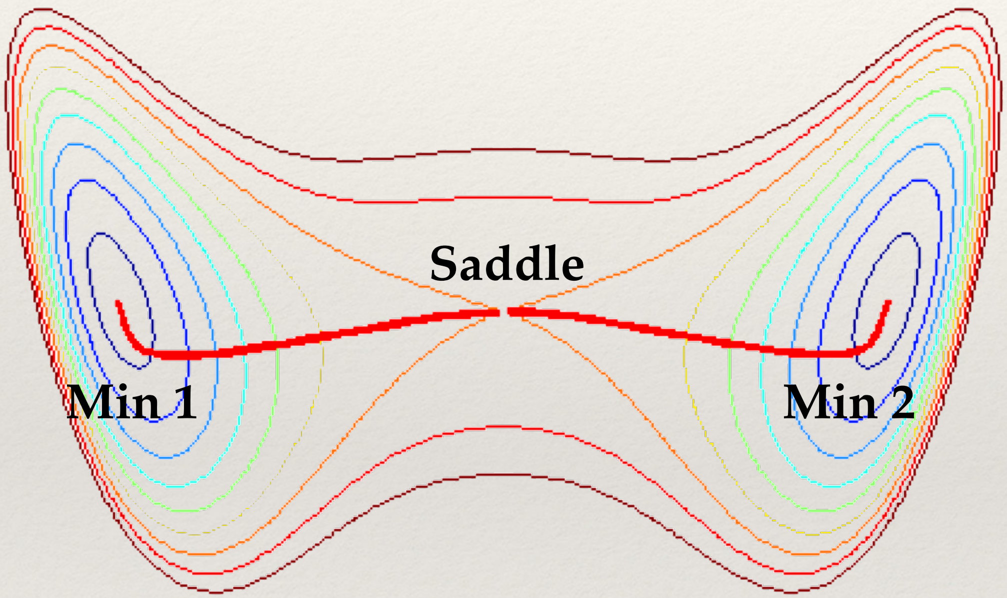




- Population dynamics
  - Dynamics of savanna landscapes (Touboul et al. 2017), 3D or 4D
  - Consumer-resource model (Collie & Spencer (1994), Steele and Henderson (1981)), 2D

Gradient SDEs

$$dx = -\nabla V(x)dt + \sqrt{2\beta^{-1}}dw$$



Arrhenius formula (1884):

$$Rate \propto e^{-\beta(V_{sad} - V_{min})}$$

Gibbs (1839–1903) measure:

$$\mu(x) = Z^{-1} e^{-\beta V(x)}$$

Kramers'/Langer's formula:
1940 1969

$$Rate \approx \frac{|\lambda|}{2\pi}\sqrt{\frac{\det H_{min}}{|\det H_{sad}|}}e^{-\beta(V_{sad} - V_{min})}$$

Large Deviation Theory

Freidlin and Wentzell, 1970s

$$dx = b(x)dt + \sqrt{\epsilon}dw$$

Freidlin-Wentzell action functional

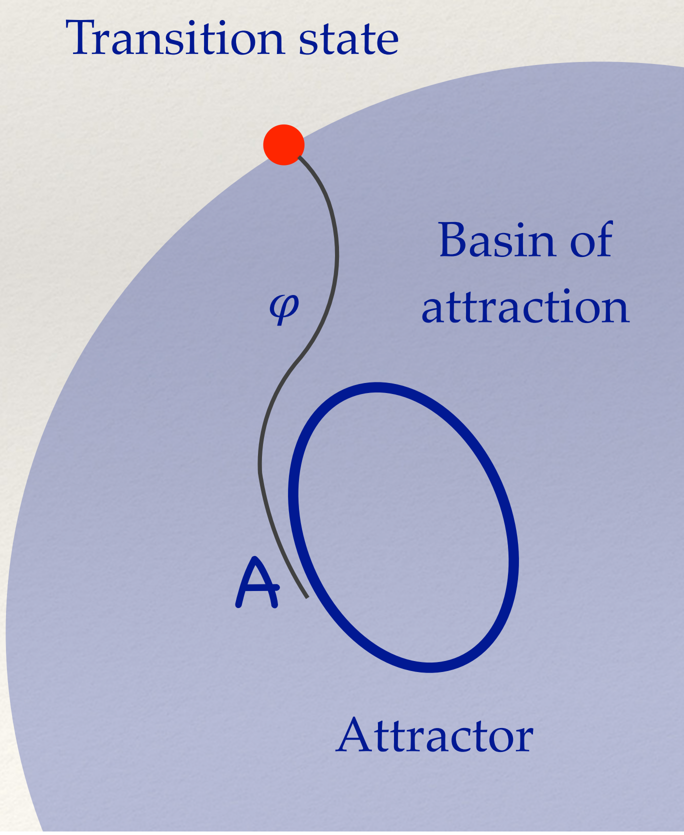
$$S_T(\phi) = \frac{1}{2}\int_0^T \|\dot{\phi} - b(\phi)\|^2 dt$$

Quasipotential

$$U_A(x) = \inf_{\phi,T}\{S_T(\phi) \mid \phi(0) \in A,\ \phi(T) = x\}$$

Expected escape time

$$\tau_A(D) \asymp \inf_{x\in\partial D} e^{U_A(x)/\epsilon}$$



Nongradient Case

Hamilton-Jacobi-Bellman

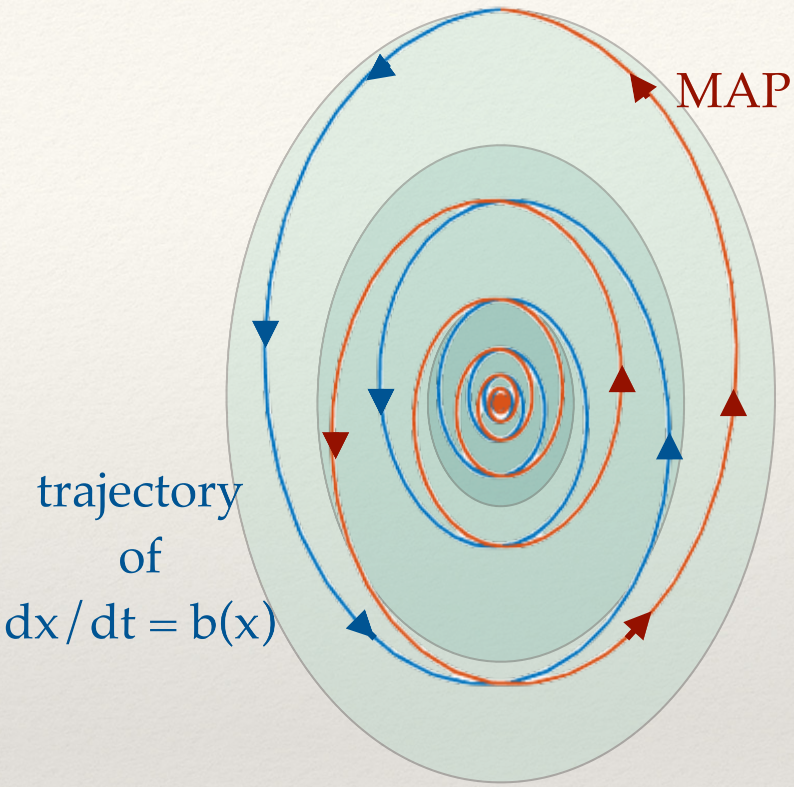
PDE for the quasi-potential

$$\frac{1}{2}\|\nabla U\|^2 + b(x)\cdot\nabla U = 0$$

$$U(x) = 0,\ \ x \in A$$

Minimum Action Paths

$$(\psi^*)' \ \parallel \ \nabla U(\psi^*) + b(\psi^*)$$



Orthogonal decomposition

$$b(x) = -\frac{1}{2}\nabla U(x) + l(x)$$

$$l(x) := \frac{1}{2}\nabla U(x) + b(x)$$

$$l(x) \ \perp \ \nabla U(x)$$

$$(\psi^*)' \ \parallel \ \frac{1}{2}\nabla U(\psi^*) + l(\psi^*)$$

$$\frac{dx}{dt} = -\frac{1}{2}\nabla U(x) + l(x)$$

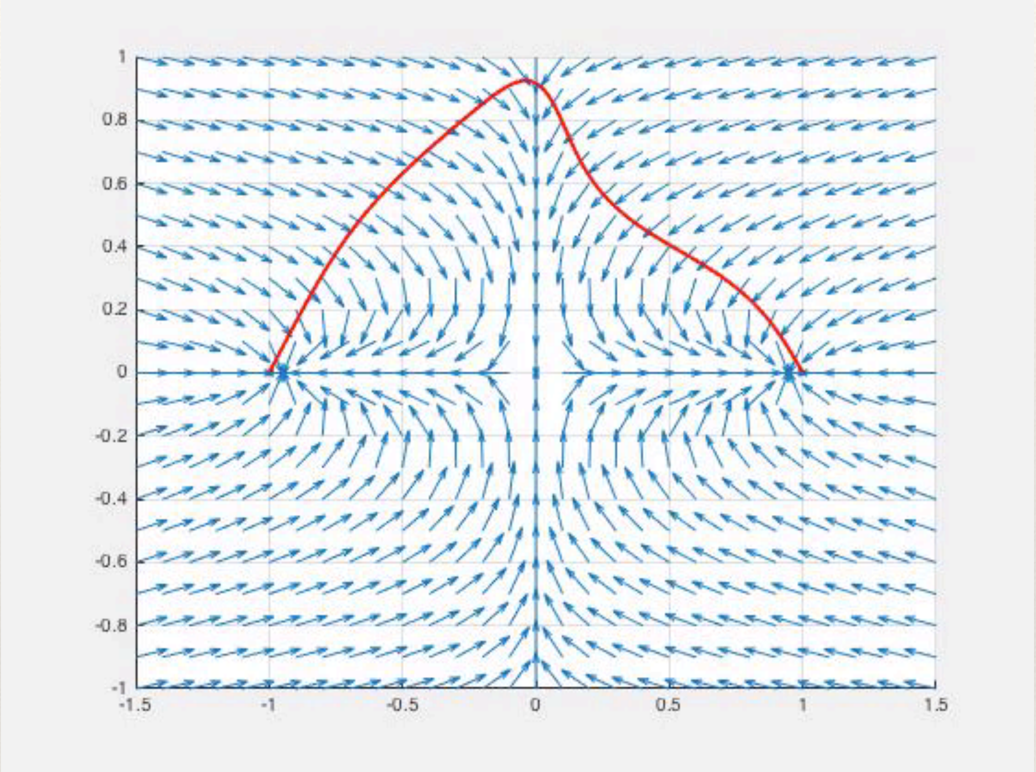
Goal: develop methods for computing the quasipotential

The quasipotential computed in a whole region allows us to estimate:

- ***expected escape times*** from the basins of attraction
- ***maximum likelihood escape paths*** (***global minimizers*** in the path space)
- ***quasi-invariant probability measures*** near attractors

Direction 1: Path-based methods

- E, Ren, Vanden-Eijnden: **MAM** (2004)
- Heymann and Vanden-Eijnden: **GMAM** (2008) (numerical minimization of the geometric action
- Zhou, Ren, and E: **AMAM** (2008) *(numerical minimization of the Freidlin-Wentzell action)*
- L. Kikuchi et al. (2020) **Ritz method** for paths and quasipotential for rare diffusive events. *Makes use of Chebyshev basis.*



Maier-Stein model, 1990s

Cameron, Dahiya, Yang, Potter, Paskal (2012,2017,2018,2019,2021)

Direction 2: computing the quasipotential

Freidlin-Wentzell action

$$S_T(\phi) = \frac{1}{2}\int_0^T \|\dot\phi - b(\phi)\|^2 dt$$

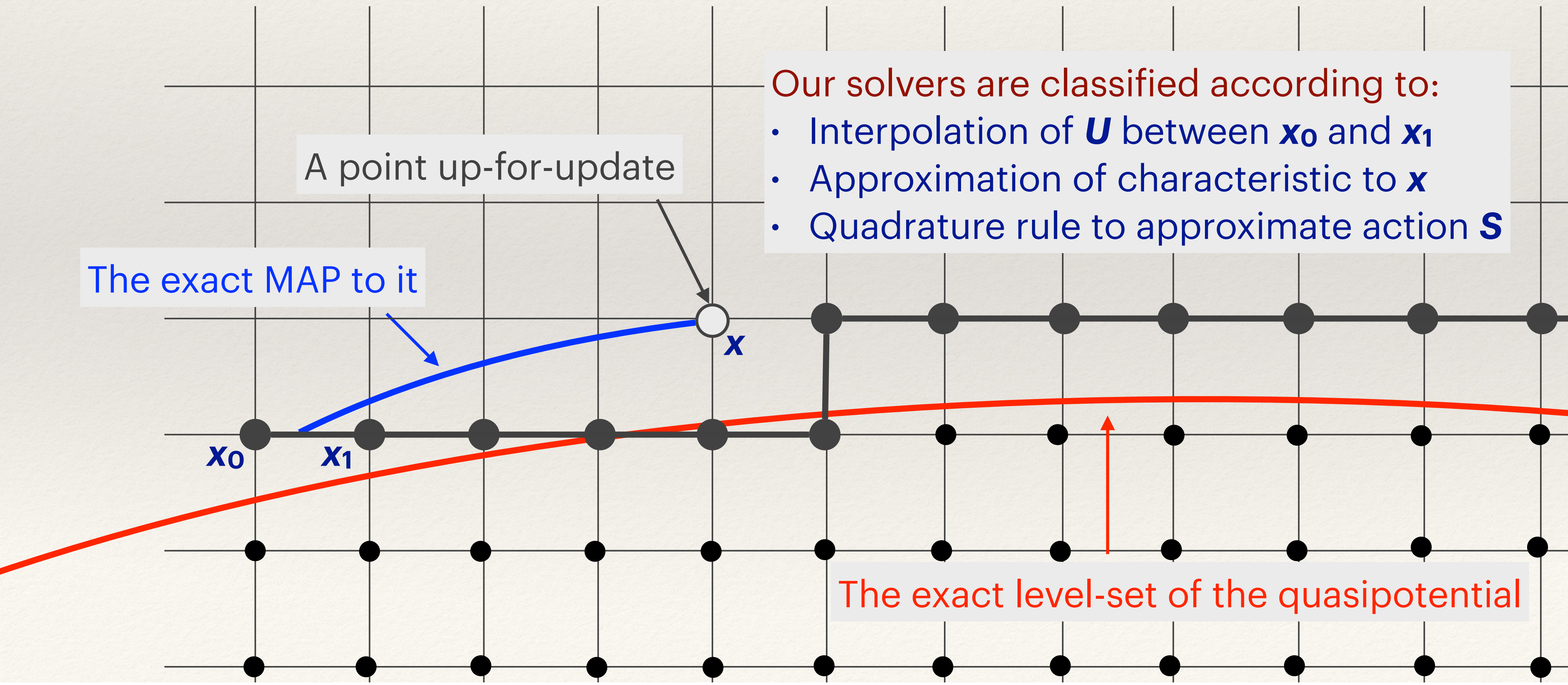
$$U_A(x) = \inf_{\phi,T}\{S_T(\phi) \mid \phi(0)\in A,\ \phi(T)=x\}$$

Analytic minimization in time

Geometric action

$$S(\psi) = \int_0^L \|\psi'\|\|b(\psi)\| - \psi'\cdot b(\psi) ds$$

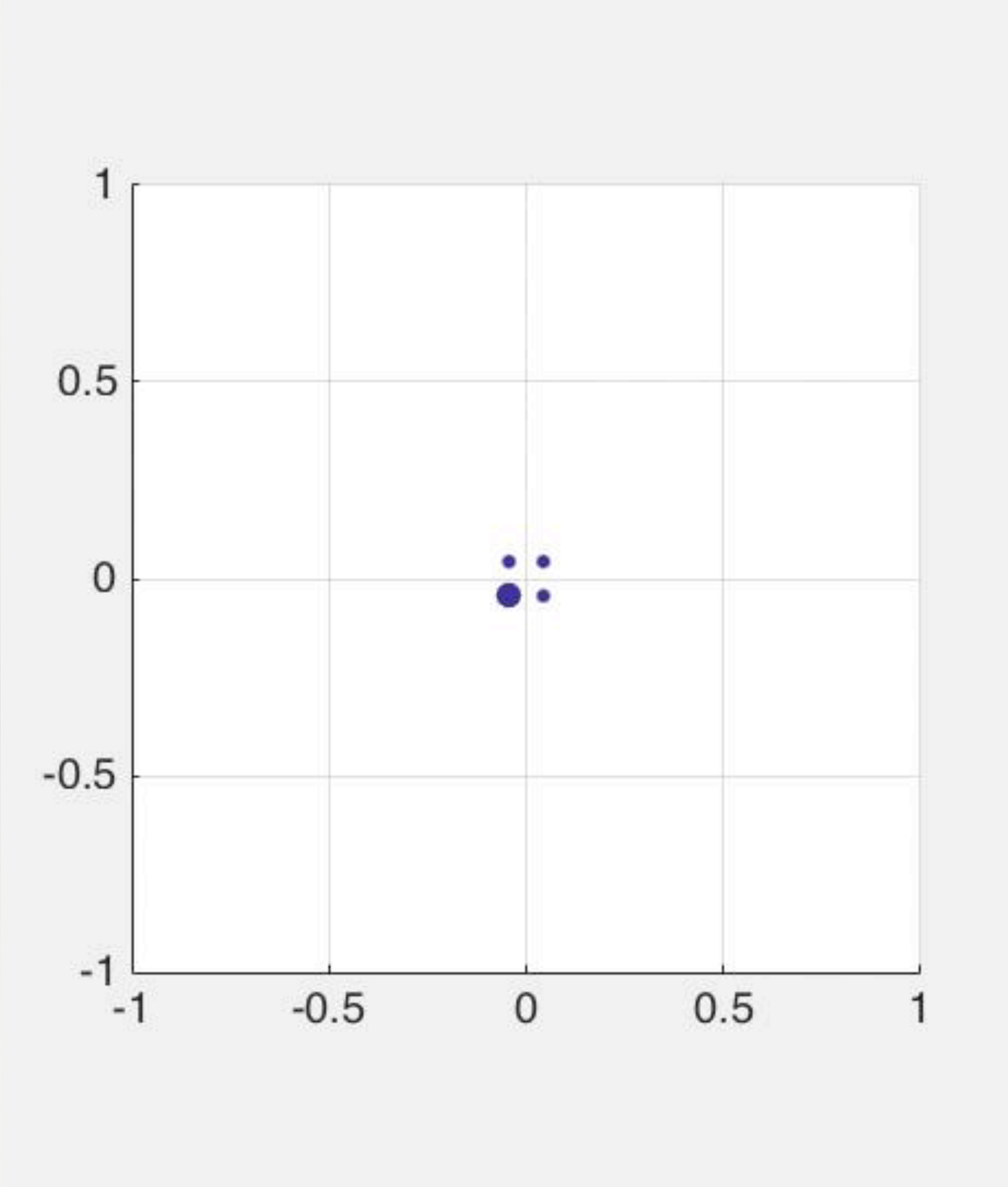
$$U(x) = \inf_{\psi}\{S(\psi) \mid \psi(0)\in A,\ \psi(L)=x\}$$



Our solvers are classified according to:

- Interpolation of U between x_0 and x_1
- Approximation of characteristic to x
- Quadrature rule to approximate action S

Ordered Line Integral Methods



4 types of mesh points:

Unknown: U is not available

Considered: U is tentative

Accepted Front:
U is finalized,
has Considered nearest neighbors

Accepted: U is finalized,
no Considered nearest neighbors

One-point and triangle updates

The minimization problem for the quasipotential

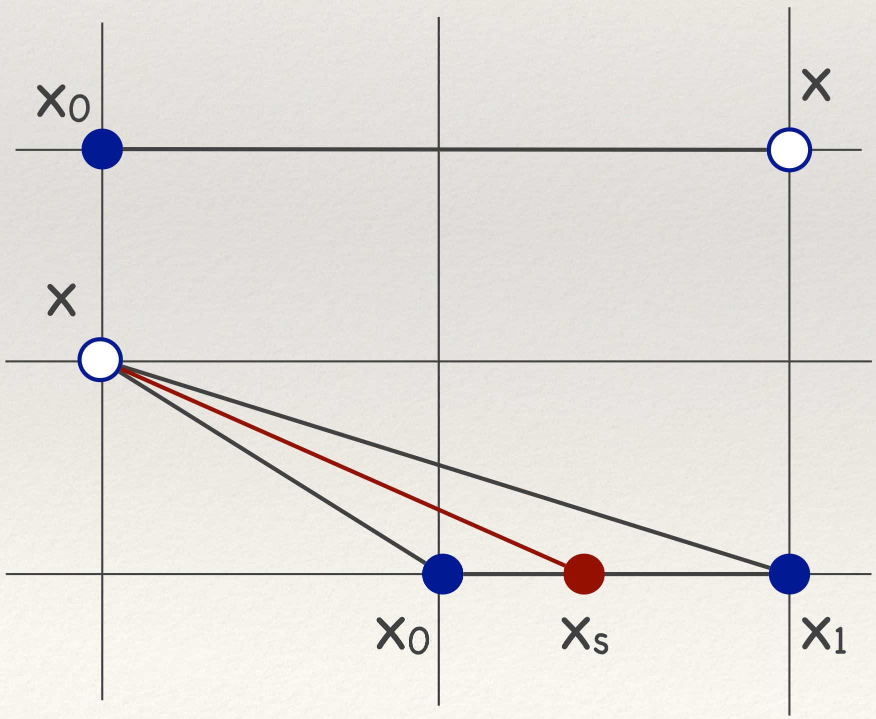
$$S(\psi) = \int_0^L \|\psi'\|\|b(\psi)\| - \psi' \cdot b(\psi) ds$$

$$U(x) = \inf_{\psi}\{S(\psi) \mid \psi(0) \in A,\ \psi(L) = x\}$$

Approximate S(ψ) with a quadrature rule

$\mathcal{Q}(a, b)$ = quadrature rule:

Right-hand: OLIM-R
Midpoint: OLIM-MID
Trapezoid: OLIM-TR
Simpson: OLIM-SIM



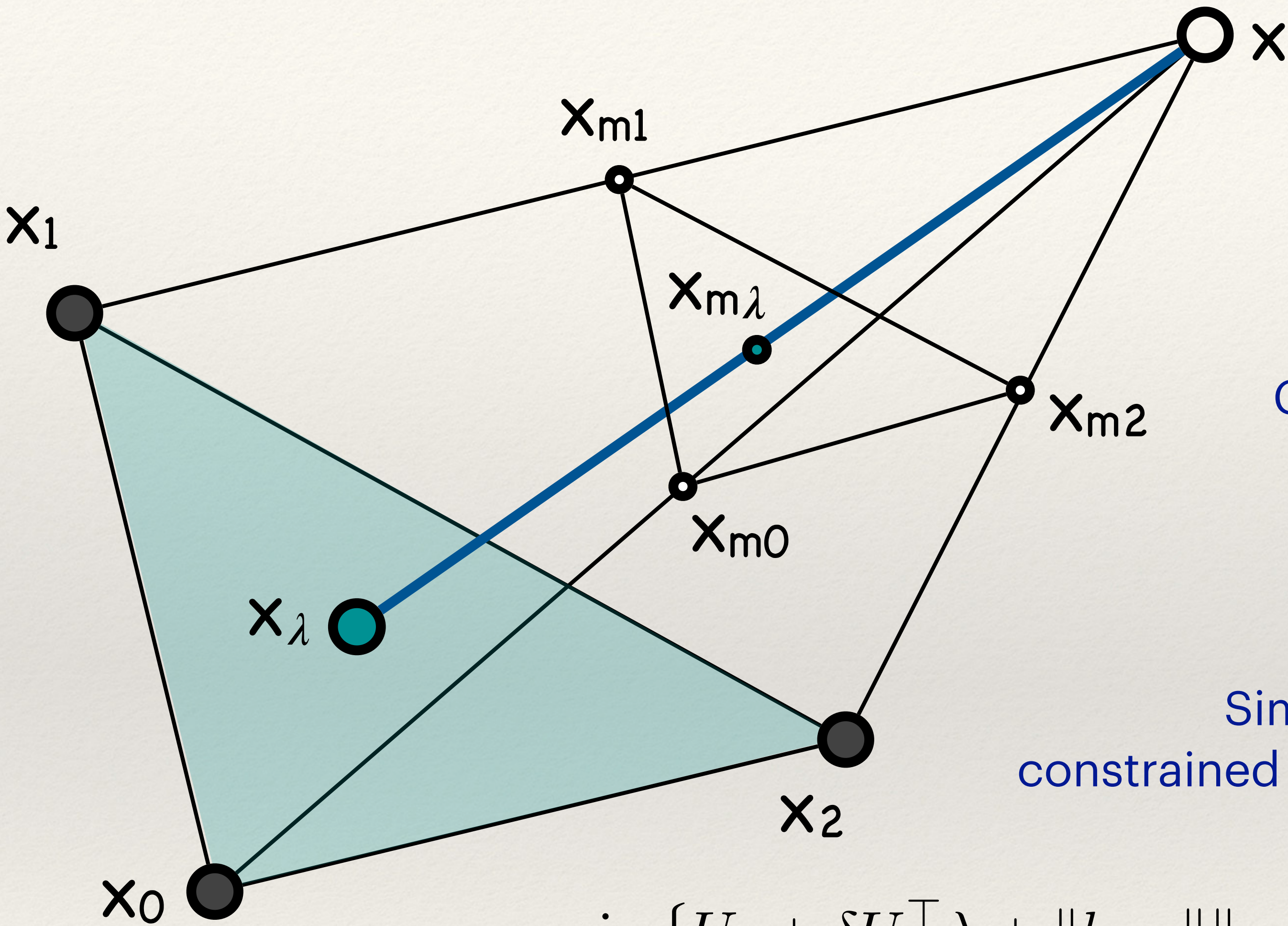
One-point update

$$\mathsf{Q}_{1pt}(x_0, x) = U(x_0) + \mathcal{Q}(x_0, x),$$
$$U(x) = \min\{\mathsf{Q}_{1pt}(x_0, x), U(x)\}.$$

Triangle update

$$\mathsf{Q}_{2pt}(x_0, x_1, x) =$$
$$\min_{s\in[0,1]} \{U_0 + s(U_1 - U_0) + \mathcal{Q}(x_s, x)\}$$
$$U(x) = \min\{\mathsf{Q}_{2pt}(x_0, x_1, x), U(x)\}$$

Simplex update



The midpoint rule
is optimal in balancing
CPU time and accuracy.

Simplex update
constrained minimization problem

$$\min \left\{ U_0 + \delta U^\top \lambda + \|b_{m\lambda}\| \|x - x_\lambda\| - b_{m\lambda}^\top (x - x_\lambda) \right\}$$

$$\text{subject to} \quad \lambda_1 \geq 0, \quad \lambda_2 \geq 0, \quad \lambda_1 + \lambda_2 \leq 1$$

Making the solver efficient

- **Hierarchical update strategy**
    - Routine 1pt update
    - Minimizer of 1pt update —> triangle update
    - Successful triangle update —> simplex update
- **Use the KKT theory** to reject simplex updates that are unlikely to succeed in finding an inner point solution
    - Starting point = minimizer of a triangle update
    - Check sign of Lagrange multiplier,
    - Reject simplex update if the KKT criteria are met
- **Restrict the set of admissible simplexes** and devise a **fast search** for them

KKT test

The problem being solved

$$\min f(\lambda) \quad \text{where} \quad f(\lambda) = U_\lambda + \|b_{m\lambda}\| \|x - x_\lambda\| - b_{m\lambda} \cdot (x - x_\lambda)$$
$$\text{subject to} \quad \lambda_1 \geq 0, \quad \lambda_2 \geq 0, \quad 1 - \lambda_1 - \lambda_2 \geq 0$$

Lagrangian function

$$L(\lambda, \mu) = f(\lambda) - \mu_1 \lambda_1 - \mu_2 \lambda_2 - \mu_3 (1 - \lambda_1 - \lambda_2)$$

The KKT necessary optimality conditions

$$\nabla_\lambda L(\lambda, \mu) = \nabla f - \mu_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \mu_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \mu_3 \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\mu_1 \geq 0, \quad \mu_2 \geq 0, \quad \mu_3 \geq 0,$$
$$\lambda_1 \geq 0, \quad \lambda_2 \geq 0, \quad 1 - \lambda_1 - \lambda_2 \geq 0,$$
$$\lambda_1 \mu_1 = 0, \quad \lambda_2 \mu_2 = 0, \quad (1 - \lambda_1 - \lambda_2)\mu_3 = 0.$$

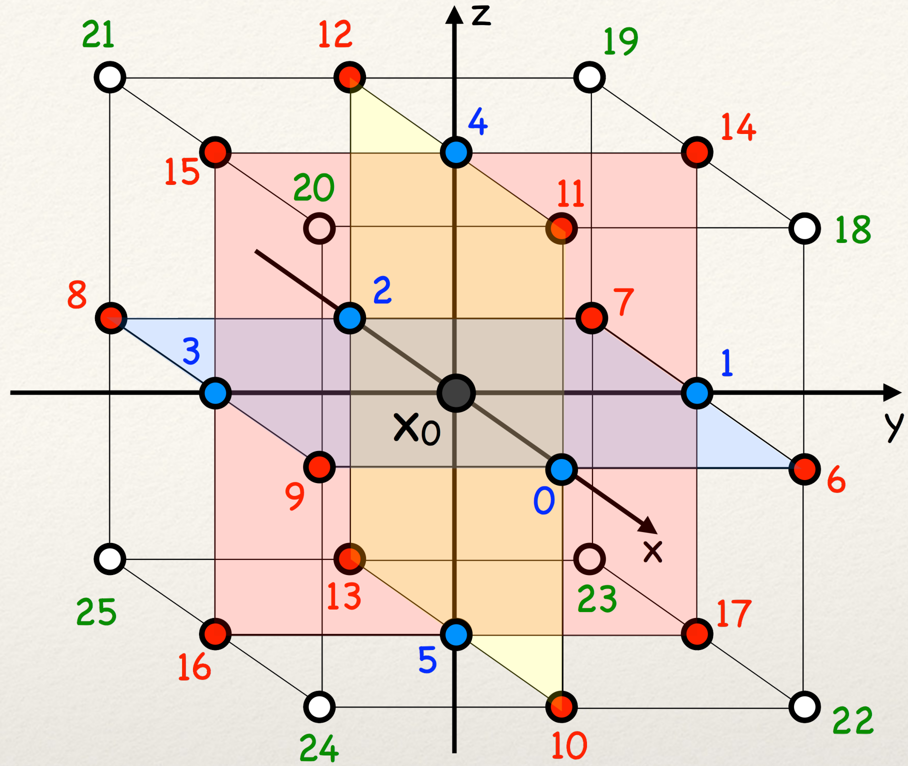
Our test: $(\lambda^*, 0)$ = initial guess.
triangle update minimizer

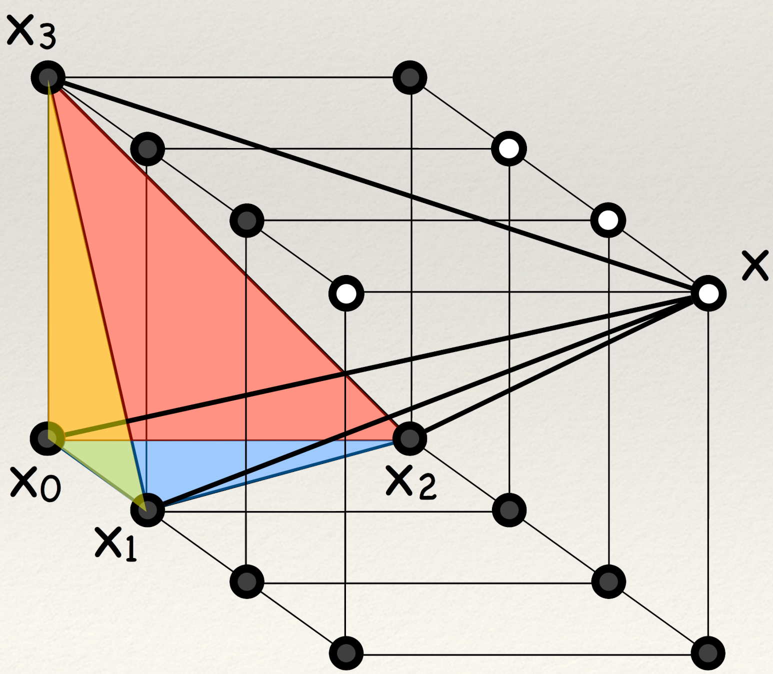
Hence: $\partial_{\lambda_1} f(\lambda^*, 0) = 0, \quad \mu_1 = 0, \quad \mu_3 = 0$

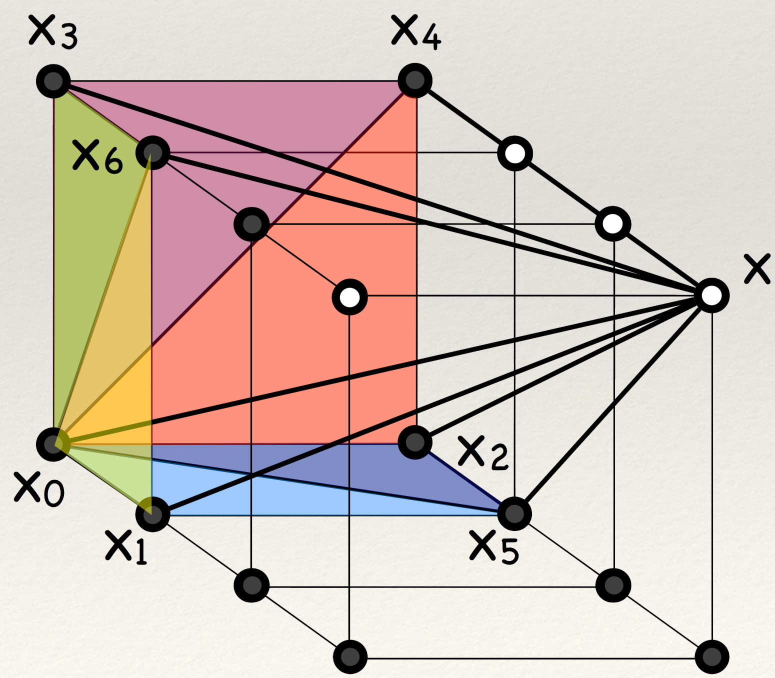
If $\partial_{\lambda_2} f(\lambda^*, 0) = \mu_2 < 0$ do simplex update

If $\partial_{\lambda_2} f(\lambda^*, 0) = \mu_2 \geq 0$ skip simplex update

- **Restrict the set of admissible simplexes** and devise a **fast search** for them







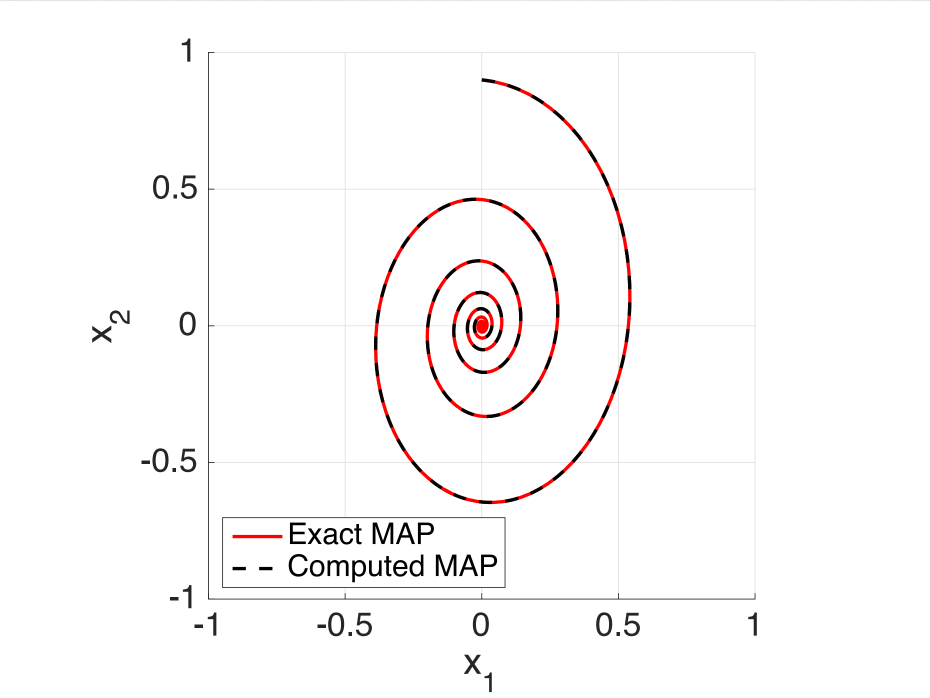
Performance tests

2D test problems

Linear

$$dx_1 = (-2x_1 - 10x_2)dt + \sqrt{\epsilon}dw_1$$
$$dx_2 = (20x_1 - x_2)dt + \sqrt{\epsilon}dw_2$$

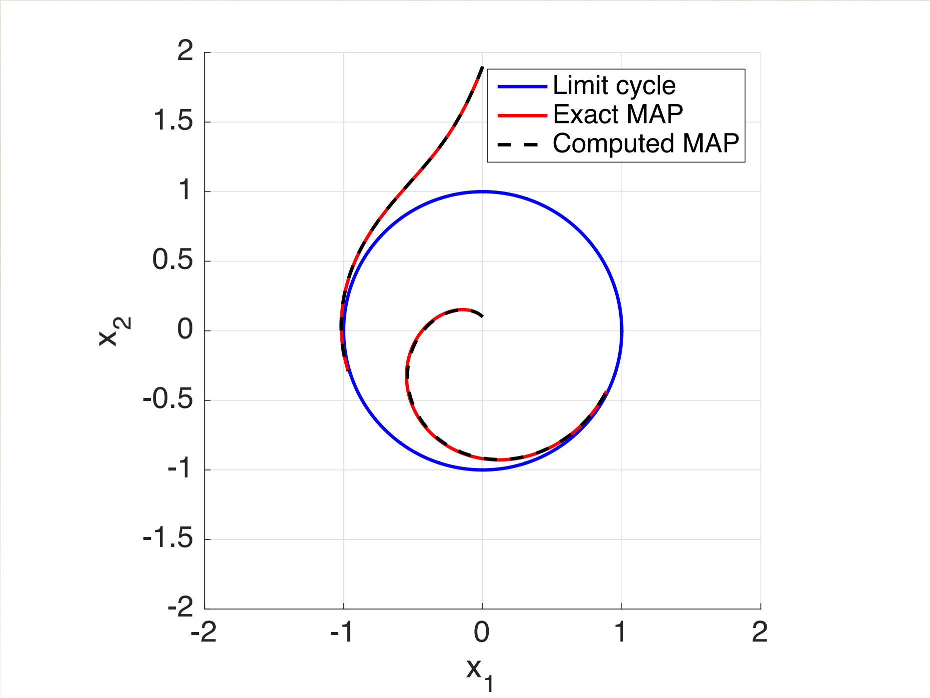
$$\boxed{U(x_1, x_2) = 2x_1^2 + x_2^2}$$



Circle

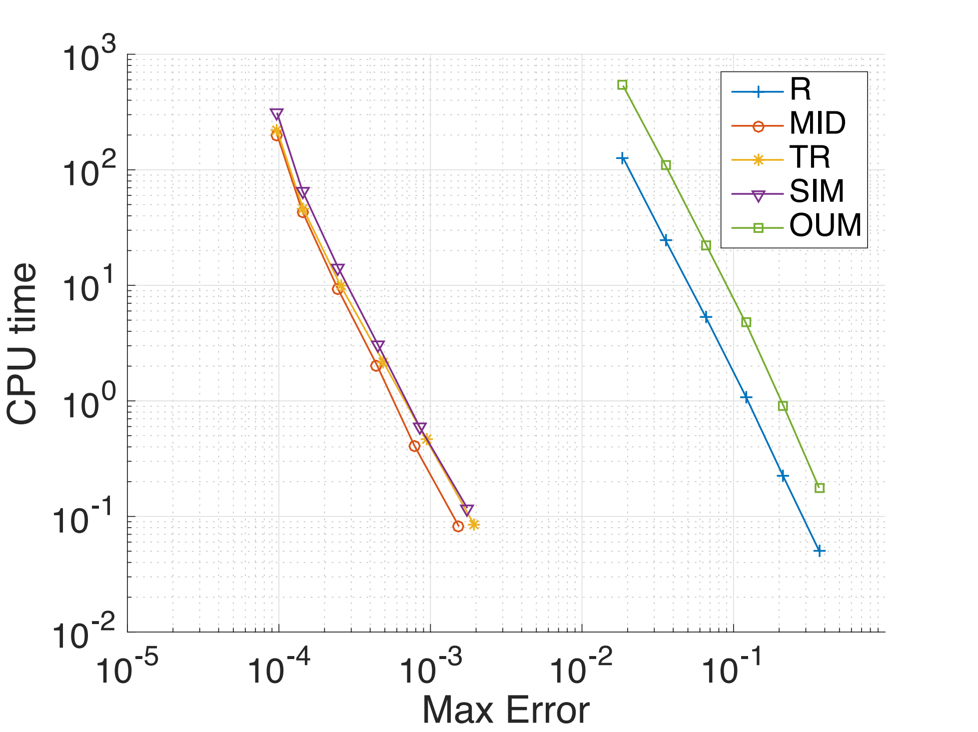
$$dx_1 = (x_2 + x_1(1 - r^2))dt + \sqrt{\epsilon}dw_1$$
$$dx_2 = (-x_1 + x_2(1 - r^2))dt + \sqrt{\epsilon}dw_2$$
$$r^2 := x_1^2 + x_2^2$$

$$\boxed{U(x_1, x_2) = \tfrac{1}{2}(x_1^2 + x_2^2 - 1)^2.}$$

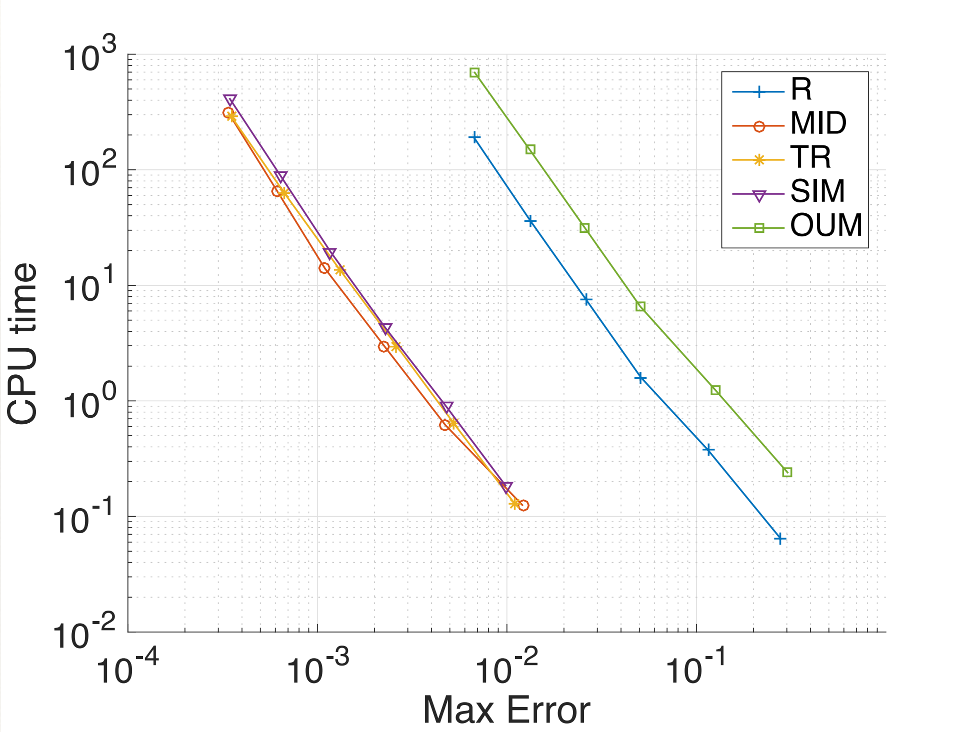


OLIMs vs OUM

Linear

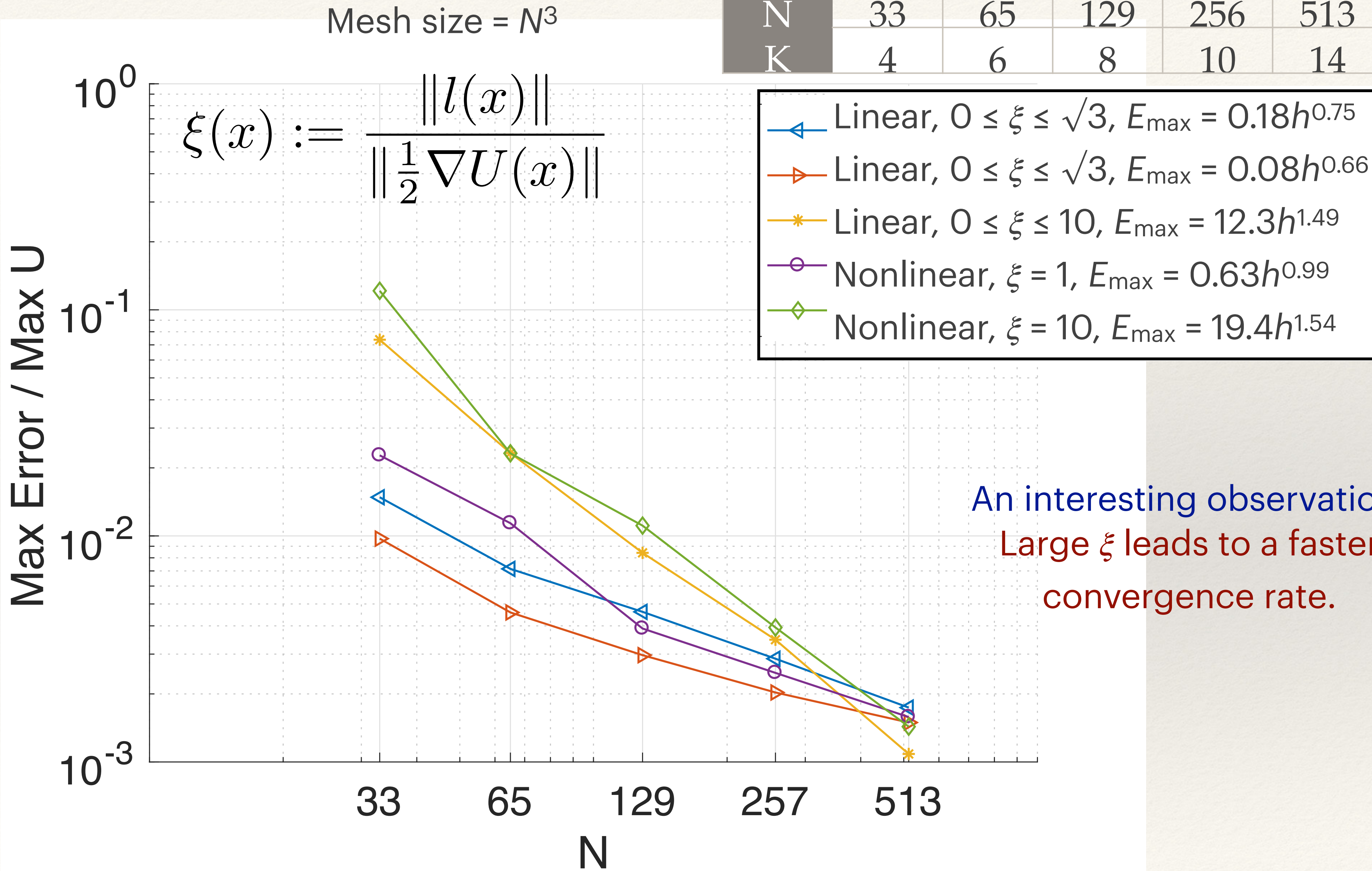


Circle



3D performance test

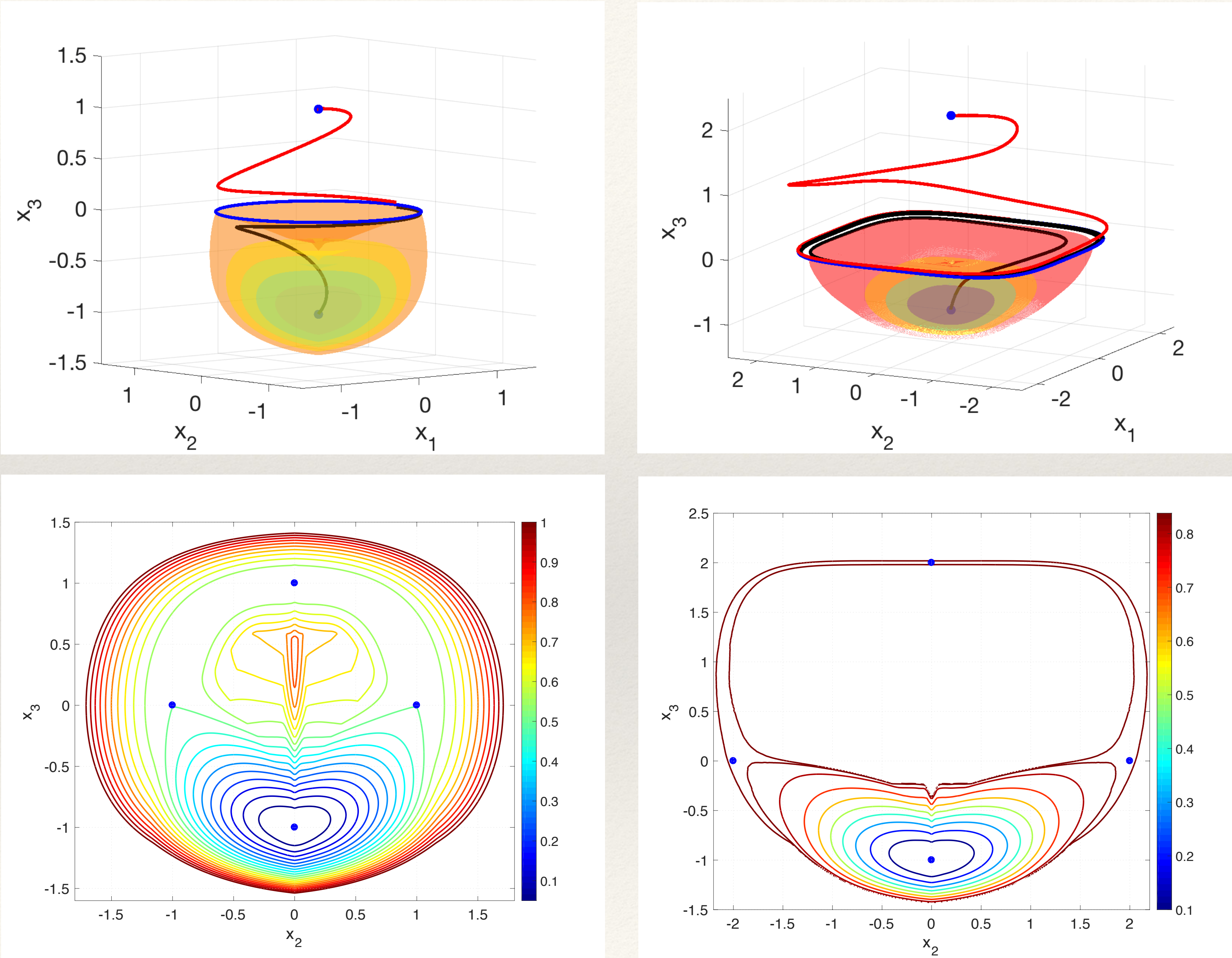
N	33	65	129	256	513
K	4	6	8	10	14



An interesting observation:
Large ξ leads to a faster
convergence rate.

Applications

3D systems with hyperbolic periodic orbits



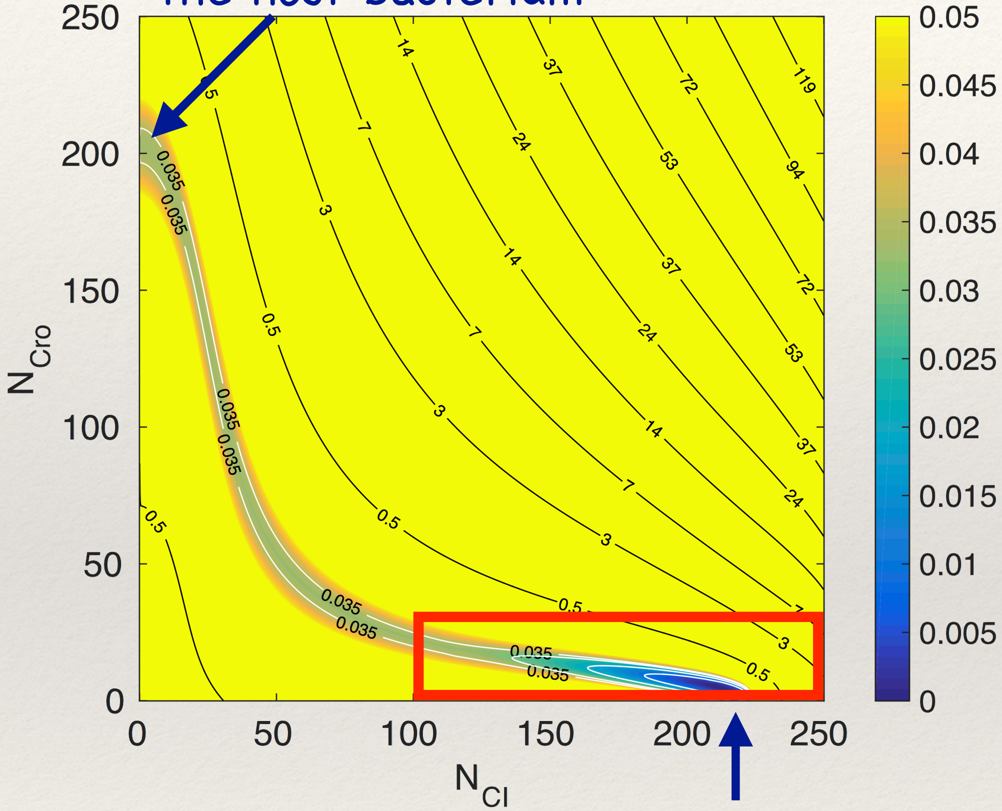
Molei Tao's examples, 2018

2D: Genetic switch in phage λ

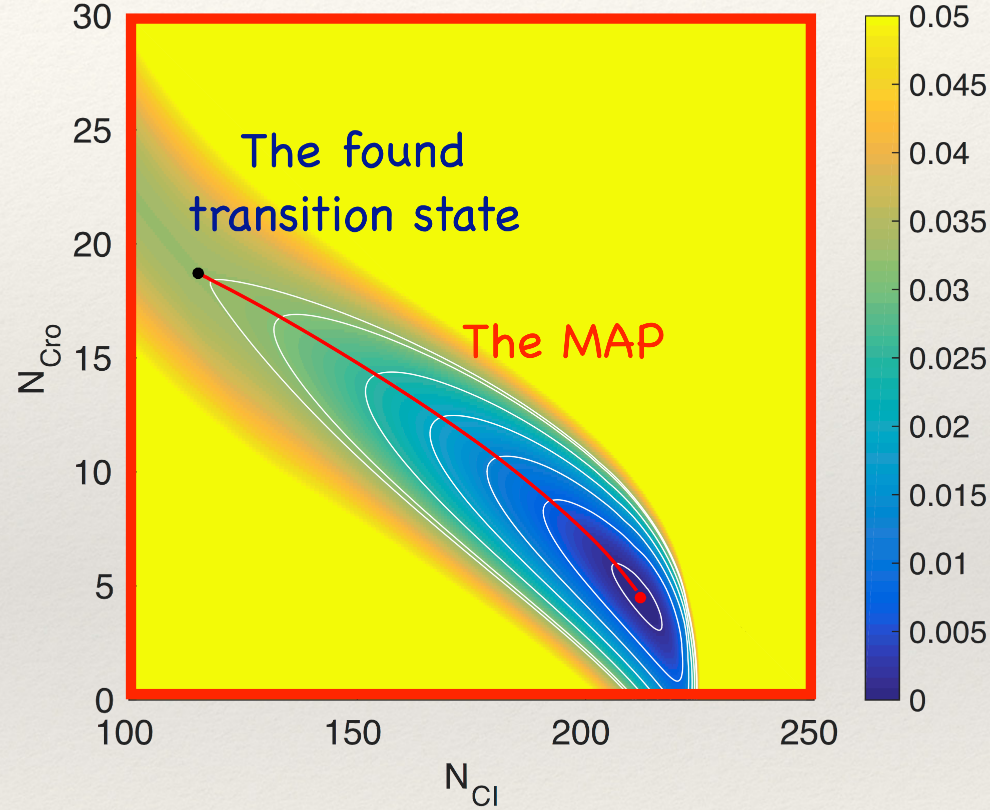
SDE model: Aurell-Sneppel, 2002

$$dX_t = b(X_t)dt + \sigma(X_t)\sqrt{\epsilon}dw_t$$

Lytic state:
phage λ kills
the host bacterium



Lysogenic state:
phage λ reproduces without
killing the host bacterium



We also found the prefactor
using Bouchet-Reygner formula

3D: Genetic switch model with a positive feedback

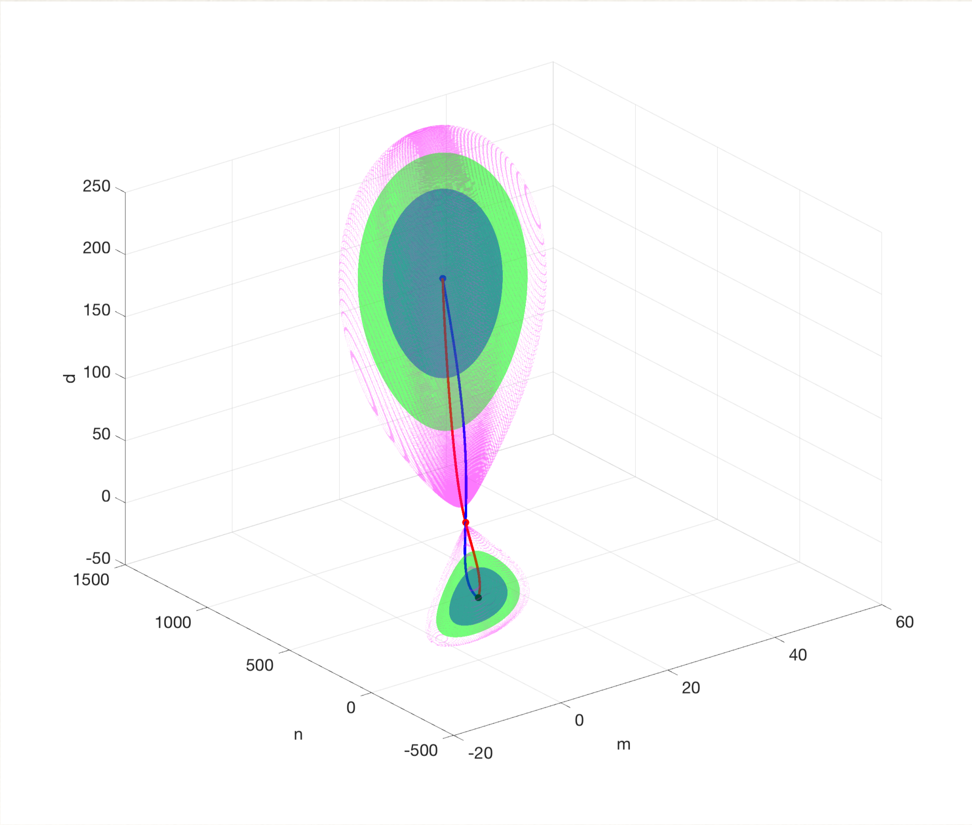
Lv, Li, Li, Li , 2014

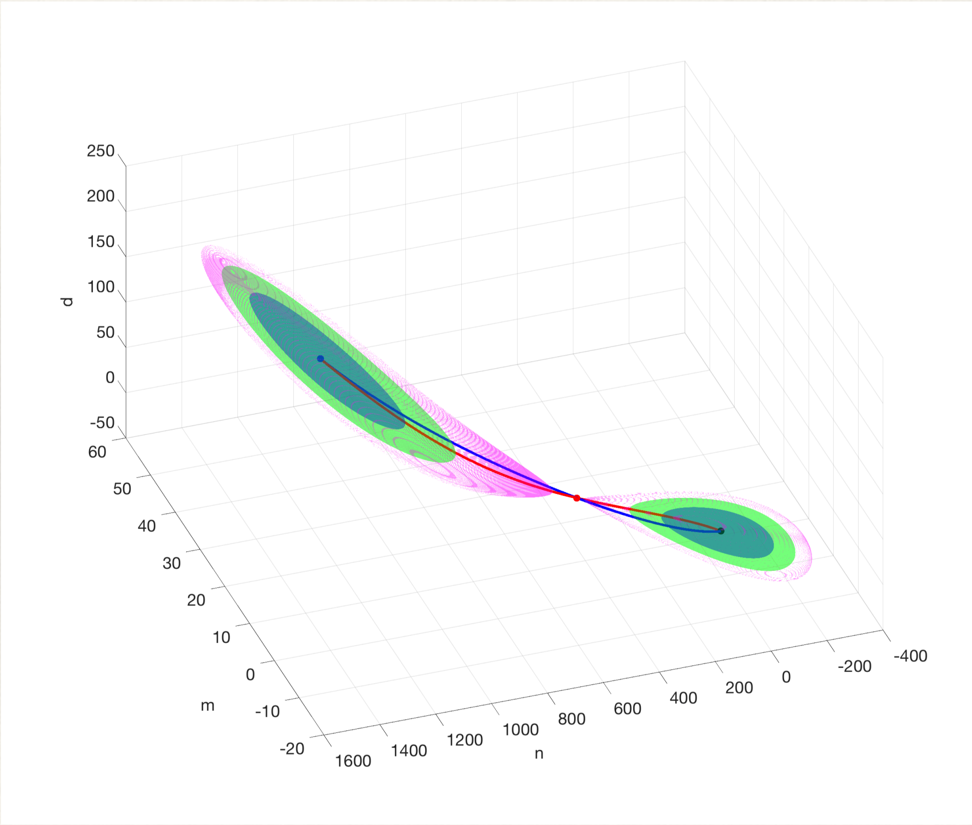
$$d\mathsf{m} = \left(\frac{a_0\gamma_0 + ak_0\mathsf{d}}{\gamma_0 + k_0\mathsf{d}} - \gamma_m\mathsf{m}\right) + \sqrt{\epsilon}dw_1,$$

$$d\mathsf{n} = \left(b\mathsf{m} - \gamma_n\mathsf{n} - 2k_1\mathsf{n}^2 + 2\gamma_1\mathsf{d}\right) + \sqrt{\epsilon}dw_2,$$

$$d\mathsf{d} = \left(k_1\mathsf{n}^2 - \gamma_1\mathsf{d}\right) + \sqrt{\epsilon}dw_3.$$

m = #mRNA, n = #protein, d = #dimer





Lorenz'63

$$\begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = \begin{bmatrix} \sigma(y-x) \\ x(\rho - z) - y \\ xy - \beta z \end{bmatrix} dt + \sqrt{\epsilon} dw$$

$$\sigma = 10, \quad \beta = 8/3, \quad 0 < \rho < \infty$$

Some critical values of ρ:

ρ = 1: the origin turned from a sink to a saddle, equilibria C_+ and C_- at $\left(\pm\sqrt{\beta(\rho-1)}, \pm\sqrt{\beta(\rho-1)}, \rho - 1\right)$ are born

$\rho \simeq$ 13.926: "preturbulence" starts (Kaplan & Yorke, 1979)

$\rho \simeq$ 24.06: the strange attractor is born (Yorke & Yorke, 1979)

$\rho \simeq$ 24.74: equilibria lose stability

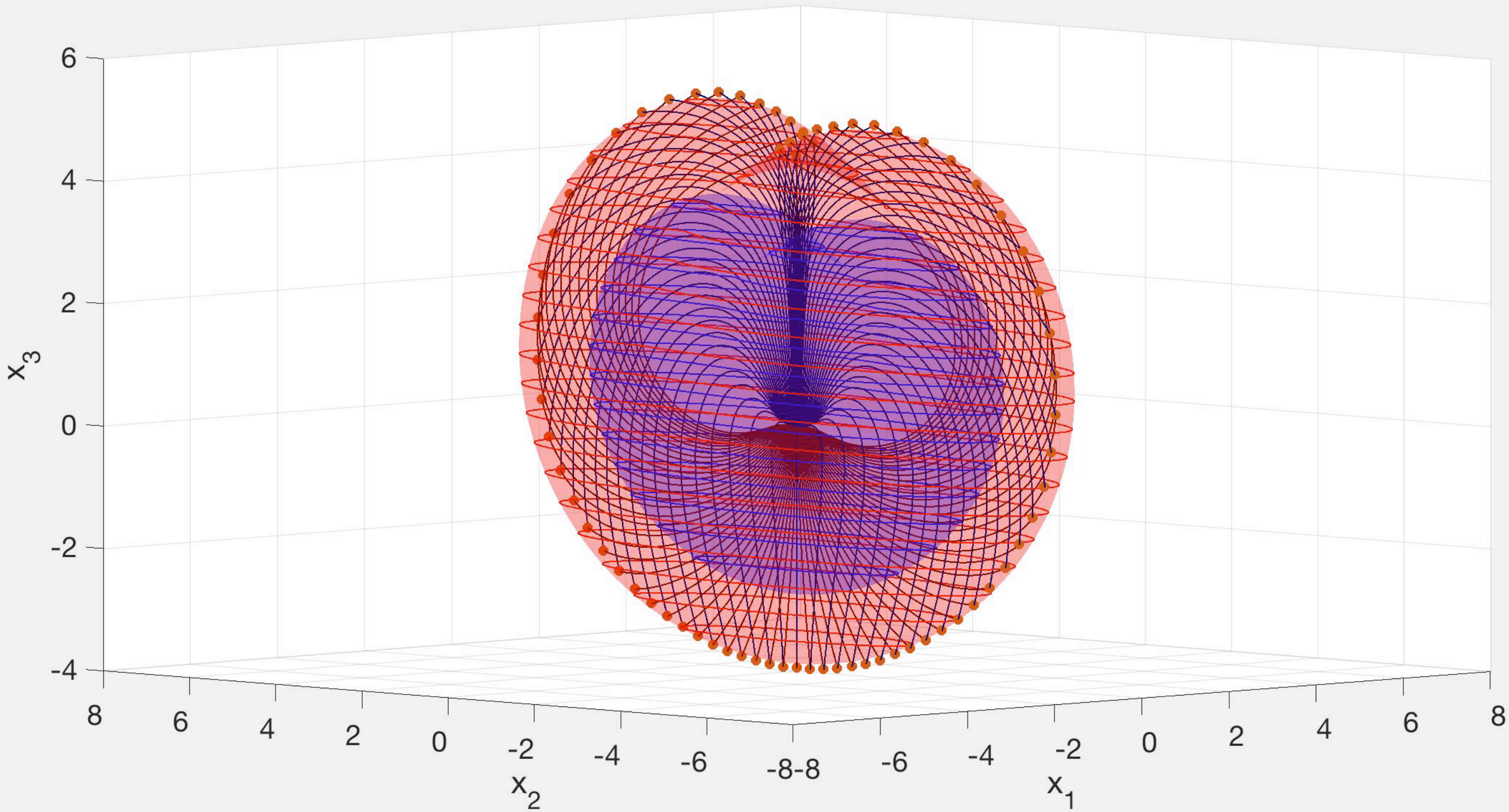
Nongradient case

Lorenz'63, $\sigma = 10$, $\beta = 8/3$, $\rho = 0.5$.

Surfaces: level sets of the quasipotential (U = 20, 40)

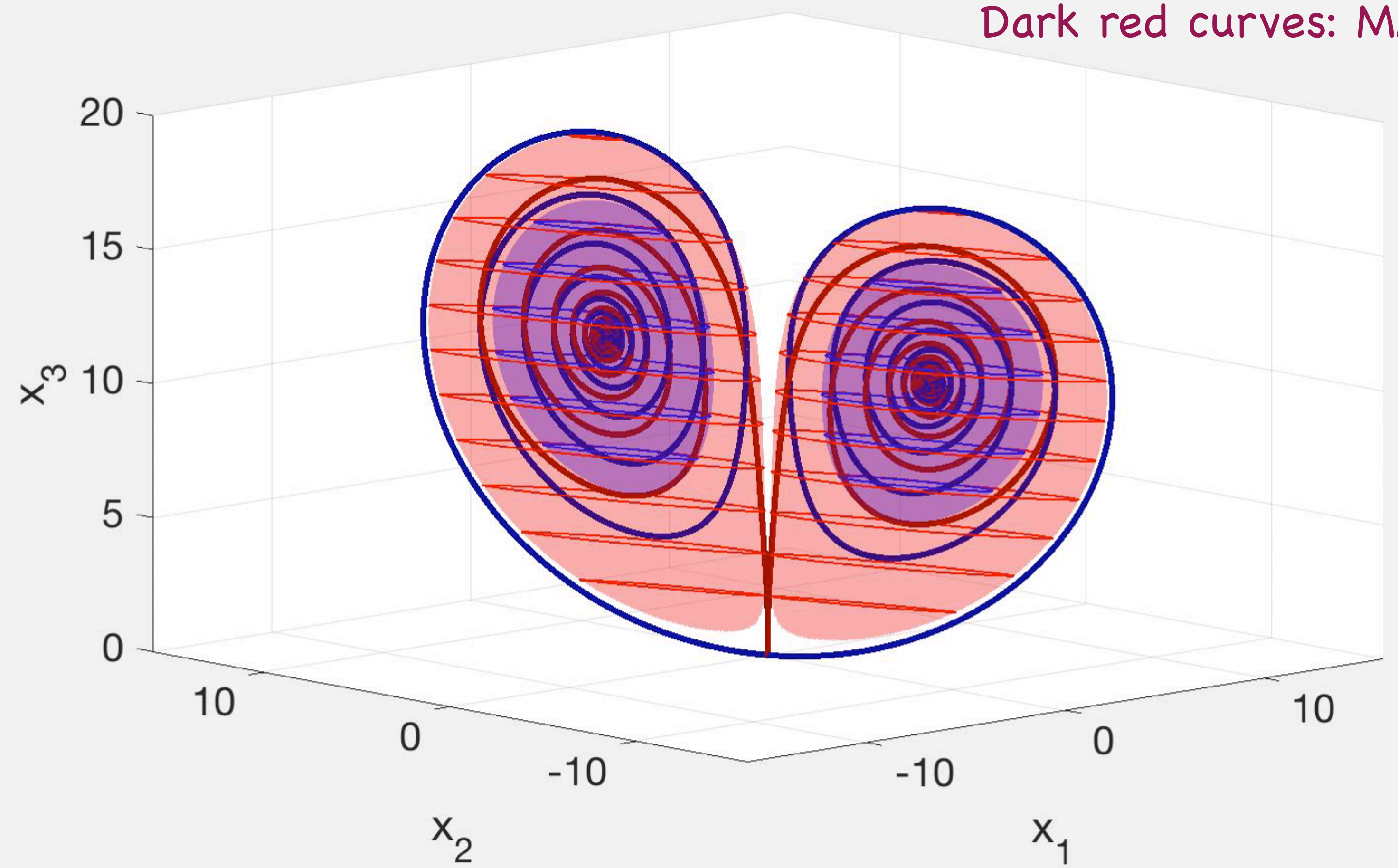
Indigo curves: trajectories

Dark red curves: MAPs

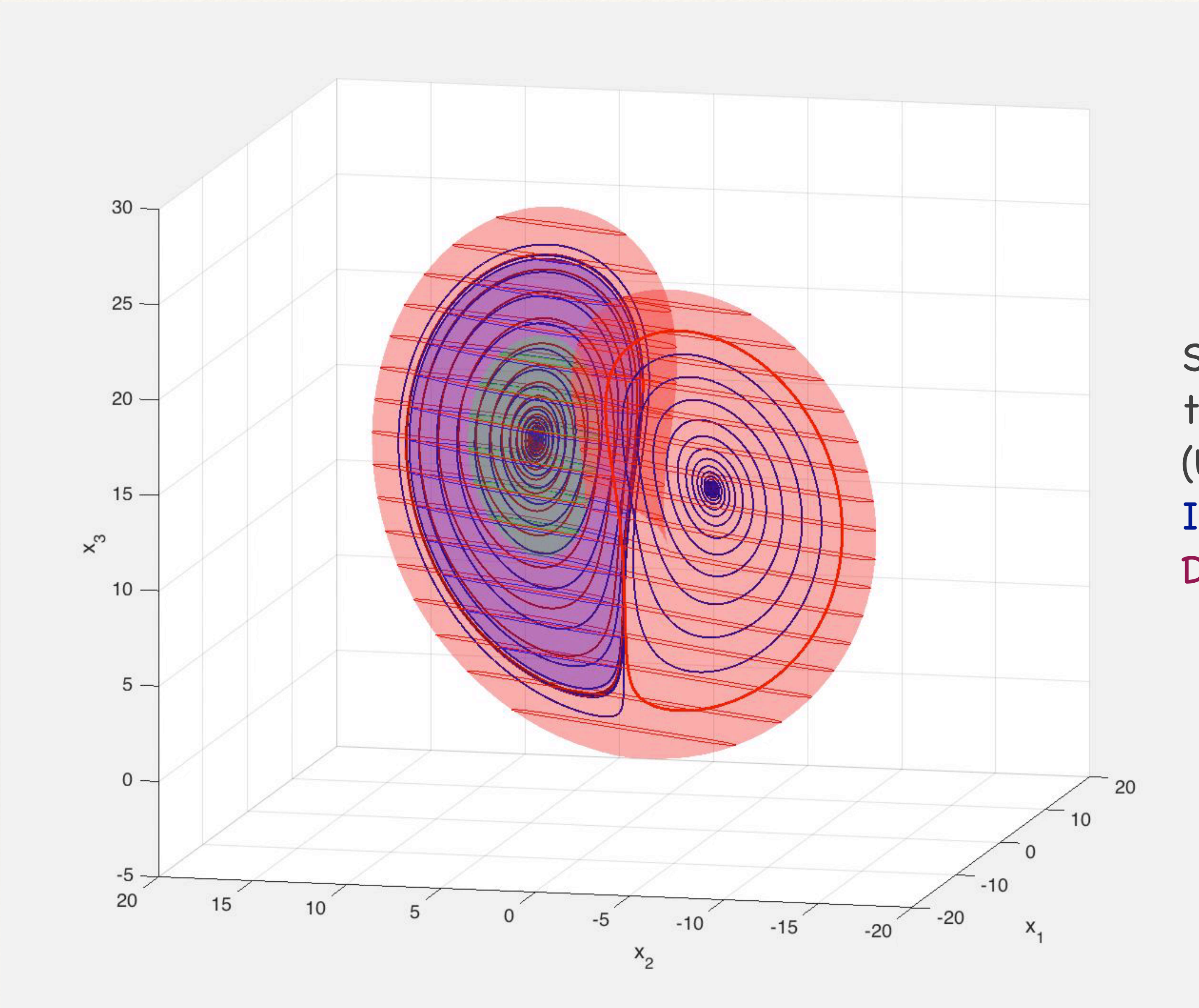


$\rho = 12$

Surfaces: level sets of the quasipotential
(U = 10, U_{origin} - 0.05)
Indigo curves: trajectories
Dark red curves: MAPs



$\rho = 15$



Surfaces: level sets of the quasipotential

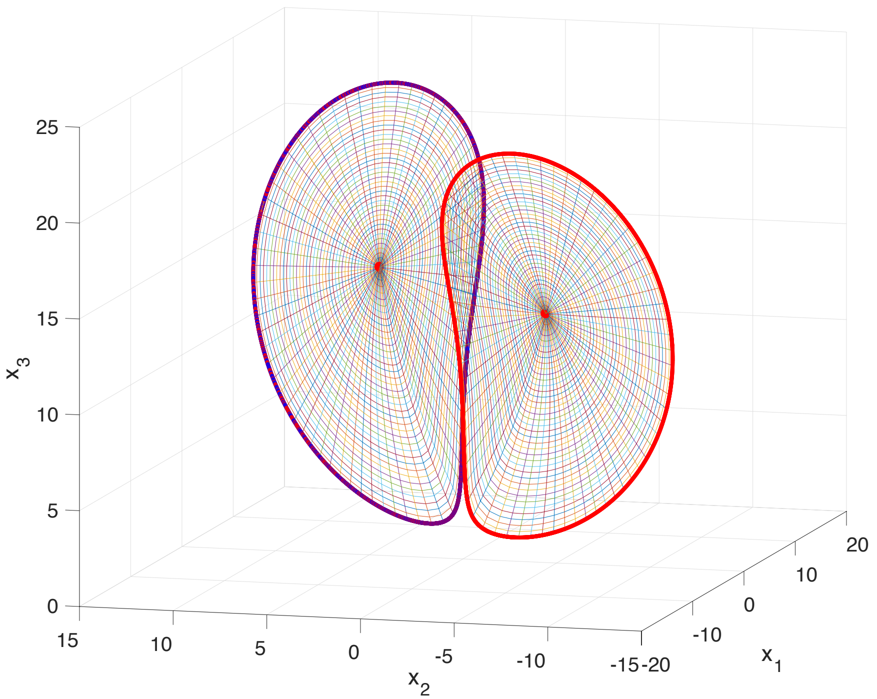
(U = 8, $U_{\text{limit cycle}}$ − 0.01, 20)

Indigo curves: trajectories

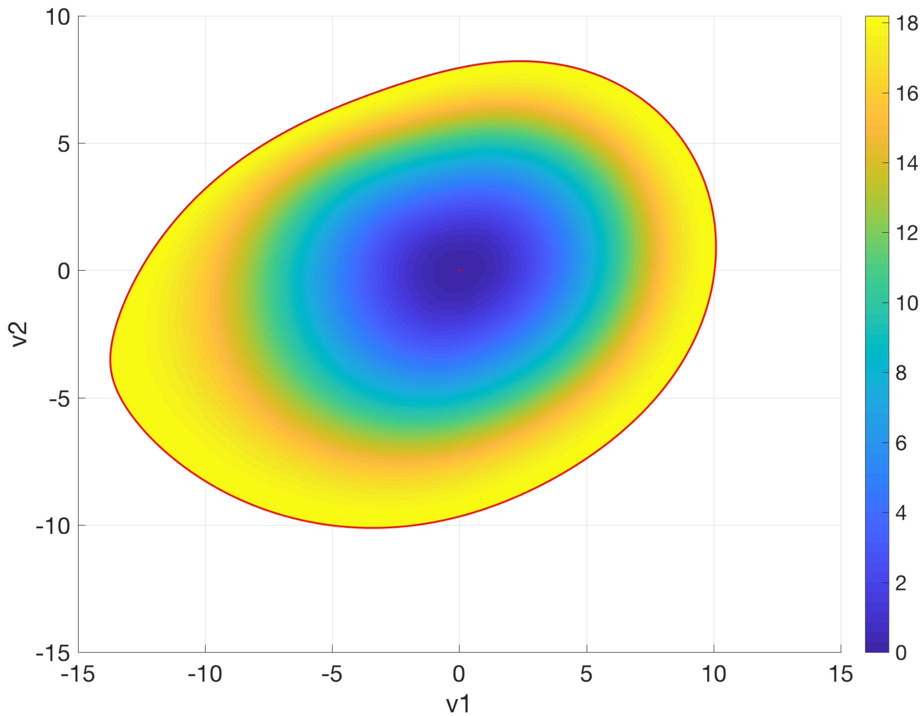
Dark red curves: MAPs

$\rho = 15$

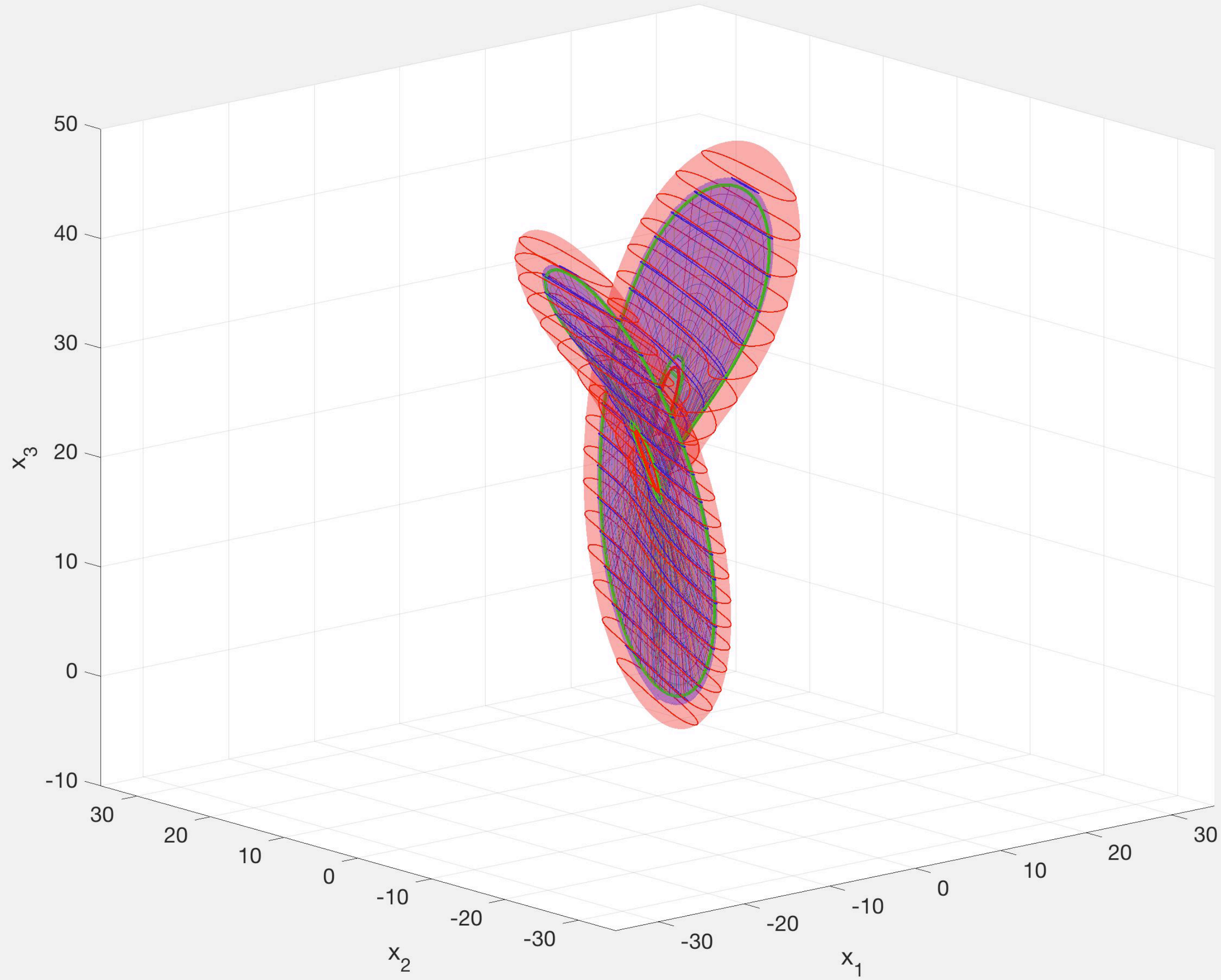
A 2D mesh on manifolds near which the dynamics are focused



The quasi potential computed on this 2D mesh



$\rho = 24.4$



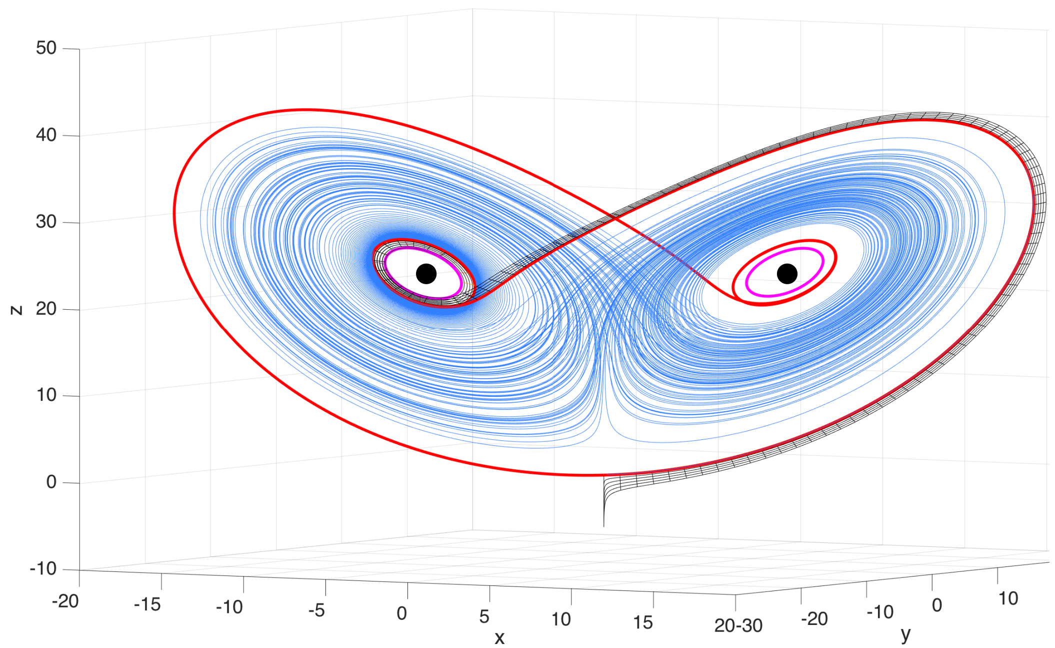
Surfaces: level sets of the quasipotential

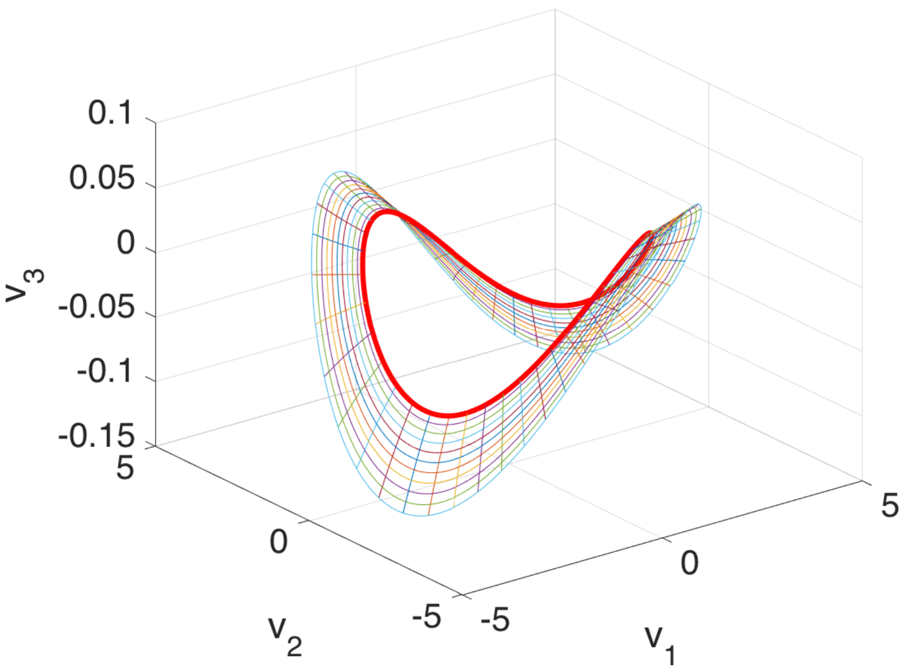
($U = U_{\text{limit cycle}} - 0.01$, 2, 20)

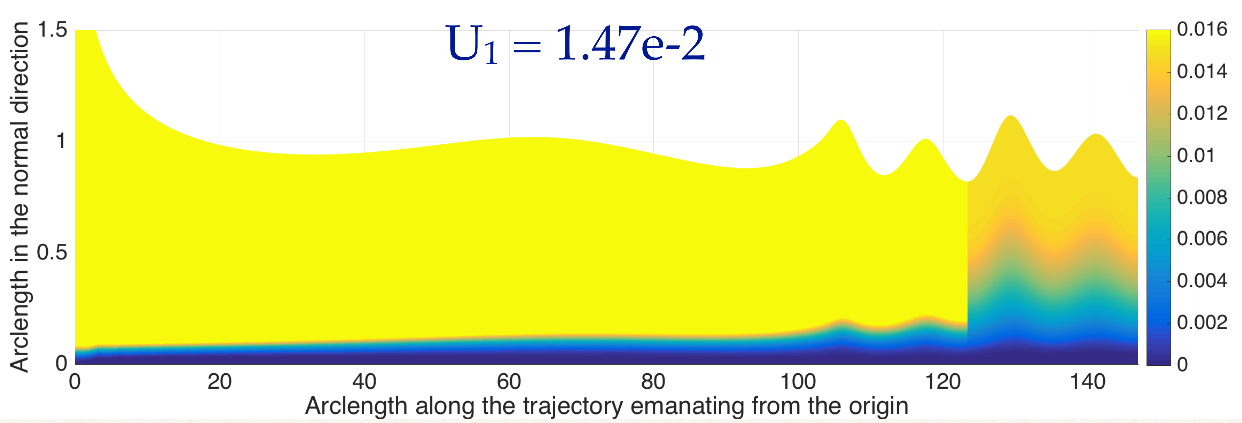
Red loops: limit cycles

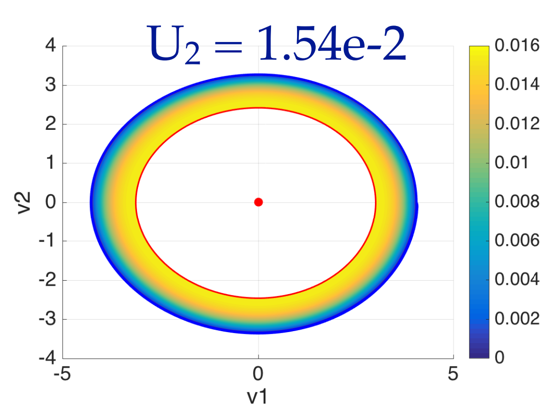
Green curves: borders of manifolds whose union approximates the strange attractor

$\rho = 24.4$: from the strange attractor to stable equilibria: which way we go?



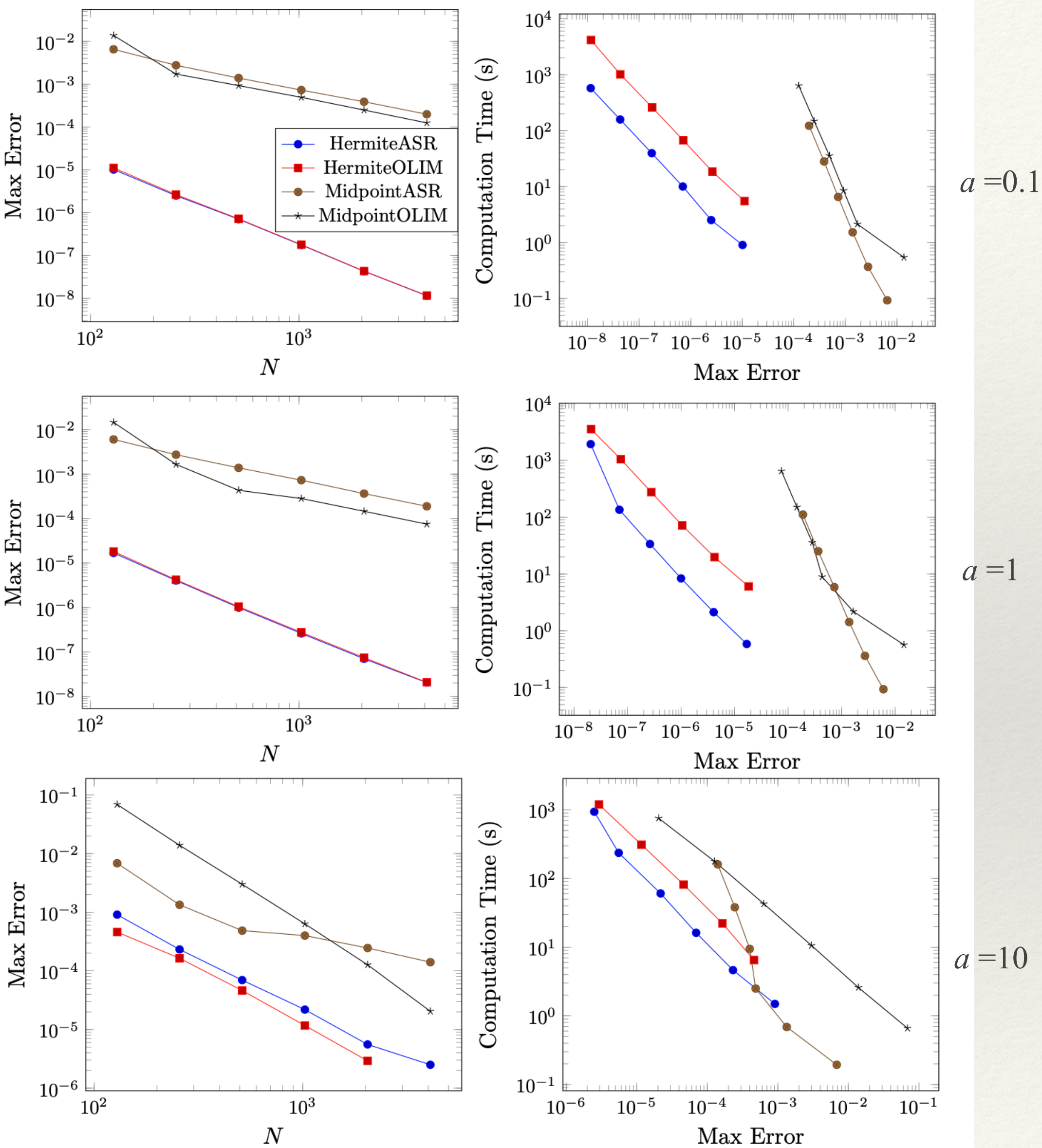






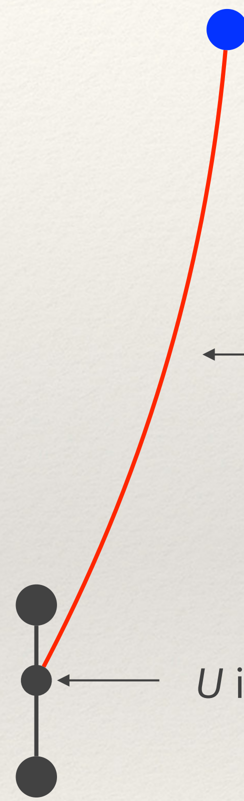
Recent advances: promotion to higher order

Three test problems



Joint work with Nick Paskal (2021)

$$b(x,y) = -\frac{1}{2}\begin{bmatrix} 4x+3x^2 \\ 2y \end{bmatrix} + \frac{a}{2}\begin{bmatrix} -2y \\ 4x+3x^2 \end{bmatrix}$$



Characteristic is approximated by a ***cubic curve***

Action is approximated by ***Simpson's rule***

U is ***Hermite***-interpolated

Search for the fastest characteristic is done by a scheme inspired by ***Mirebeau's adaptive stencil refinement***

Open problems and perspectives

- **Mechanical systems with external forcing and small noise**

$$\ddot{x} + \gamma \dot{x} + \nabla V(x) - F(t) + \sqrt{\epsilon}\eta = 0$$

- Going to **higher dimensions** by making use of methods coming from data science and machine learning (**geometric methods (diffusion maps) and neural networks**)

References

- Dahiya & Cameron, Ordered line integral methods for computing the quasipotential, J. Sci. Comp., 75/3, 1351—1384 (2018), arXiv: 1706.07509
- Dahiya & Cameron, An ordered line integral method for computing the quasipotential in the case of variable and anisotropic diffusion, Physica D 382—383, (2018) 33—45, arXiv: 1806.05321
- Yang, Potter, & Cameron, Computing the quasipotential for nongradient SDEs in 3D, J. Comp. Phys. 379, 325—350 (2019), arXiv: 1808.00562
- Cameron & Yang, Computing the quasipotential for highly dissipative and chaotic SDEs. An application to Lorenz'63, *CAMCoS* 14-2(2019), 207—246, arXiv:1809.09987v2

- **C codes and user's guides:** http://www.math.umd.edu/~mariakc/index.html