# Lect 1. Averaging methods for SPEs

## A. Multiscale analysis: Concepts.

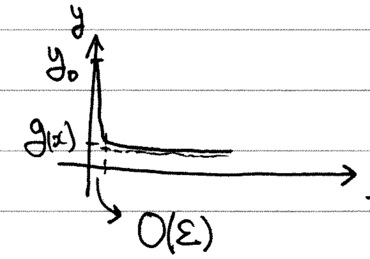
### 1. Model problem 1:

$$\begin{cases} \dfrac{dx}{dt} = f(x,y) \\ \dfrac{dy}{dt} = \dfrac{1}{\varepsilon}(g(x)-y) \end{cases} \qquad \varepsilon \ll 1,\ \varepsilon > 0$$

$\begin{cases} x: \text{ slow variable} & \text{varying in } O(1) \text{ timescale} \\ y: \text{ fast variable} & \text{〃 in } O(\varepsilon) \text{ timescale} \end{cases}$

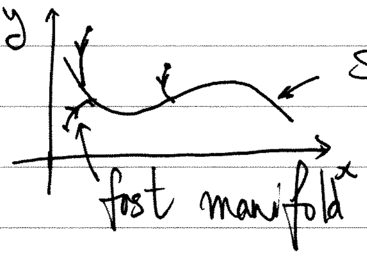
Exact solution of $y$ given $x$:

$$y(t) = e^{-t/\varepsilon} y_0 + (1 - e^{-t/\varepsilon}) g(x) \xrightarrow{t \to +\infty} g(x)$$

[ Dynamics of $y$ given $x$ :  ]



Fast relaxation to $y = g(x)$ in $O(\varepsilon)$ timescale





$\begin{cases} x: \text{ Master variable} \\ y: \text{ Slave variable} \end{cases} \qquad y = g(x): \text{ slow manifold}$

Final reduced eqn: $\frac{dx}{dt} = f(x, g(x))$, $\varepsilon \to 0$

(Effective dynamics)

Adiabatic elimination: valid in $O(\varepsilon)$ approximation.

2. Model problem 2:

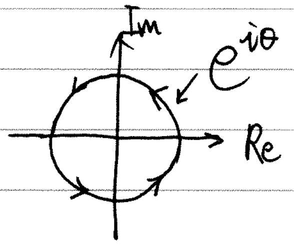
$$\dot{x} = \frac{i}{\varepsilon}x + f(x)x, \quad \varepsilon \ll 1, \quad f \in R$$

Take $x = re^{i\theta}$.

$$\dot{r}e^{i\theta} + ire^{i\theta}\dot{\theta} = \frac{i}{\varepsilon}re^{i\theta} + f(re^{i\theta})re^{i\theta}$$

$$\begin{cases} \dot{r} = f(re^{i\theta}) \triangleq g(r, e^{i\theta}) \\ \dot{\theta} = 1/\varepsilon \end{cases} \qquad \theta(t) = \theta_0 + t/\varepsilon$$

$\begin{cases} r: \text{ slow variable} \\ \theta: \text{ fast variable} \end{cases}$ $\theta(t)$ does NOT converge to a single point, how to do averaging?




$e^{i\theta}$: induces a uniform distribution on $S^1$.

$$\dot{R} = \langle g(R, z) \rangle_{\mu_R(z)} \triangleq \int g(R, z) \underbrace{\mu_R(dz)}_{\mathcal{U}(S^1)}$$

$$= \frac{1}{2\pi}\int_0^{2\pi} g(R, e^{i\theta})\, d\theta$$

3. Model ~~problem~~ problem 3:

$$\begin{cases} \dfrac{dx}{dt} = f(x,y) \\ \dfrac{dy}{dt} = \dfrac{1}{\varepsilon}(g(x)-y) + \sqrt{\dfrac{2}{\varepsilon}}\,\dot{W} \end{cases} \quad \left(\text{or. } dY_t = \frac{1}{\varepsilon}(g(X_t)-Y_t)dt + \sqrt{\frac{2}{\varepsilon}}\,dW_t\right)$$

$\dot{W}_t$ = White noise in time (Gaussian) with $\mathbb{E}\dot{W}_t = 0$, $\mathbb{E}(\dot{W}_t\dot{W}_s) = \delta(t-s)$.

Given $x$, $y(t)$ has an invariant distribution:

$$y(t) \overset{t\to+\infty}{\sim} N(g(x), 1) \triangleq \mu_{g(x)}(y)dy$$

Effective dynamics:

$$\frac{dx}{dt} = \langle f(x,y) \rangle_{\mu_{g(x)}(y)} \triangleq \int_{\mathbb{R}} f(x,y)\,\mu_{g(x)}(y)dy$$

B. Analytical approach for averaging: (singular perturbation analysis)

Refs: 1. G. Papanicolaou, Introduction to the asymptotic analysis of stochastic equations, Lectures in Applied Math. 16 (1977), 109-147.

2. G. Papanicolaou, D. Stroock and S. Varadhan, Martingale approach to some limit theorems, In "Statistical Mechanics, Dynamical systems and the Duke Turbulence Conf.", Duke University Series III, 1977.

1. Framework: (semigroup approach)

$$u^{\varepsilon}(x,t) \triangleq \mathbb{E}^{x} f(X_t^{\varepsilon}) \qquad f \in C_0^{\infty}(\mathbb{R}^d), \; X_t^{\varepsilon}: \text{multiscale process}$$

Assume:

$$\begin{cases} \partial_t u^\varepsilon = \frac{1}{\varepsilon^2}\mathcal{L}_1 u^\varepsilon + \frac{1}{\varepsilon}\mathcal{L}_2 u^\varepsilon + \mathcal{L}_3 u^\varepsilon \\ u^\varepsilon\big|_{t=0} = f(x), \qquad t \in [0,T] \end{cases}$$

i) $\mathcal{L}_1$ is an infinitesimal generator:

Assume the semigroup $S(t) = e^{\mathcal{L}_1 t} \xrightarrow{t\to\infty} Q$, then

$Q^2 = Q$ is a projection operator to the null space of $\mathcal{L}_1$

NB: $Q$ need not be orthogonal projection since $Q^* \neq Q$ in general.

Example: $\mathcal{L}_1 = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix}$, $\quad N(\mathcal{L}_1) = \text{span}\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$

$$e^{\mathcal{L}_1 t} \longrightarrow Q = \begin{pmatrix} 2/3 & 1/3 \\ 2/3 & 1/3 \end{pmatrix}, \qquad \mathcal{L}_1 Q = Q\mathcal{L}_1 = 0$$

and

$$R(Q) = N(\mathcal{L}_1), \quad N(Q) = \overline{R(\mathcal{L}_1)}$$

[ $N(\mathcal{L}_1)$: Null space, $R(Q)$: Range ]

ii) Solvability condition: $Q\mathcal{L}_2 Q = 0$

iii) Consistency condition: $Qf = f$ (no initial layer)

Asymptotics:

$$u^\varepsilon = u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + \cdots$$

$O(\varepsilon^{-2})$: $\quad \mathcal{L}_1 u_0 = 0 \qquad$ ①

$O(\varepsilon^{-1})$: $\quad \mathcal{L}_1 u_1 = -\mathcal{L}_2 u_0 \qquad$ ②

$O(1)$: $\quad \mathcal{L}_1 u_2 = -\mathcal{L}_2 u_1 - \mathcal{L}_3 u_0 + \partial_t u_0 \qquad$ ③

$$u_0|_{t=0} = f(x)$$

① $\Rightarrow u_0 \in N(\mathcal{L}_1) = R(Q)$, $\quad Qu_0 = u_0$

From ②: We assume Fredholm alternative for $\mathcal{L}_1$ (which should be rigorously proved for each concrete problem)

$$\left.\begin{aligned} &\text{Condition ii)} \Rightarrow \mathcal{L}_2 u_0 \in N(Q) \\ &\forall g \in N(\mathcal{L}_1^*) \Rightarrow g \in {}^{\perp}\overline{R(\mathcal{L}_1)} = {}^{\perp}N(Q) \end{aligned}\right\} \Rightarrow \langle g, \mathcal{L}_2 u_0 \rangle = 0$$

$\exists$ solution of ②, denoted as $\quad u_1 = -\mathcal{L}_1^{-1}\mathcal{L}_2 Q u_0$

③ $\Rightarrow$

$$\begin{cases} \partial_t u_0 = (Q\mathcal{L}_3 Q - Q\mathcal{L}_2\mathcal{L}_1^{-1}\mathcal{L}_2 Q)u_0 \triangleq \bar{\mathcal{L}} u_0 \\ u_0|_{t=0} = f(x) \qquad\qquad u_0 \in R(Q) \end{cases} \qquad ④$$

Effective dynamics for $u_0$, $\bar{\mathcal{L}}$: effective operator.

<u>NB</u>: Condition ii) ensures that the choice of solution of ② does not affect the final system ④.

2. A simple example:

$Y_t = y(t)$: 2-state Markov jump process taking values $\pm\alpha$ with jumping rate $\beta$.

Generator:

$$A = \begin{pmatrix} -\beta & \beta \\ \beta & -\beta \end{pmatrix}$$

Consider the process $(x^\varepsilon(t), y^\varepsilon(t))$ defined as

$$\begin{cases} \dfrac{dx^\varepsilon}{dt} = \dfrac{1}{\varepsilon} y^\varepsilon(t), & x^\varepsilon(0) = x \\ y^\varepsilon(t) = y(t/\varepsilon^2) & \text{diffusive scaling} \end{cases}$$

Def:

$$u^\varepsilon(x,y,t) \triangleq \mathbb{E}^{x,y} f(x_t^\varepsilon, y_t^\varepsilon)$$

$$\partial_t u^\varepsilon = \frac{1}{\varepsilon} y \frac{\partial u^\varepsilon}{\partial x} + \frac{1}{\varepsilon^2} A u^\varepsilon, \qquad u^\varepsilon\big|_{t=0} = f(x,y)$$

Def: $u_\pm^\varepsilon(x,t) = u^\varepsilon(x, \pm\alpha, t)$, $f_\pm(x) = f(x, \pm\alpha)$

$$\partial_t \begin{pmatrix} u_+^\varepsilon \\ u_-^\varepsilon \end{pmatrix} = \frac{1}{\varepsilon} \begin{pmatrix} \alpha & 0 \\ 0 & -\alpha \end{pmatrix} \partial_x \begin{pmatrix} u_+^\varepsilon \\ u_-^\varepsilon \end{pmatrix} + \frac{1}{\varepsilon^2} \underbrace{\begin{pmatrix} -\beta & \beta \\ \beta & -\beta \end{pmatrix}}_{= A} \begin{pmatrix} u_+^\varepsilon \\ u_-^\varepsilon \end{pmatrix}$$

Let $w(x) = u_+^\varepsilon + u_-^\varepsilon$

$$\varepsilon^2 \partial_t^2 w = \alpha^2 \partial_x^2 w - 2\beta \partial_t w, \qquad w\big|_{t=0} = f_+ + f_-$$

$$\partial_t w\big|_{t=0} = \frac{\alpha}{\varepsilon} \partial_x (f_+ - f_-)$$

$$w = w_0 + \varepsilon w_1 + \varepsilon^2 w_2 + \cdots$$

$$\Rightarrow \frac{\partial w_0}{\partial t} = \frac{\alpha^2}{2\beta} \frac{\partial^2 w_0}{\partial x^2}, \qquad w_0\big|_{t=0} = 2f \quad (\text{Assume } f = f_+ = f_-)$$

To leading order, $x_t^\varepsilon \sim$ B.M. with diffusion constant $D = \alpha/\sqrt{\beta}$.

Applying framework:

$$\mathcal{L}_1 = A, \quad \mathcal{L}_2 = \begin{pmatrix} \alpha & 0 \\ 0 & -\alpha \end{pmatrix} \partial_x, \quad \mathcal{L}_3 = 0$$

$$Q = \lim_{t\to\infty} e^{\mathcal{L}_1 t} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

$$\mathcal{L}_1^{-1} = -\int_0^{+\infty} (e^{\mathcal{L}_1 t} - Q)\, dt = -\frac{1}{4\beta}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

$$Q\mathcal{L}_2 Q = 0, \qquad Q\vec{f} = \vec{f} \Rightarrow f_+ = f_- \triangleq f$$

Effective operator for $x$:

$$\bar{\mathcal{L}} = -Q\mathcal{L}_2\mathcal{L}_1^{-1}\mathcal{L}_2 Q = \frac{\alpha^2}{4\beta}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\frac{\partial^2}{\partial x^2}$$

Finally:

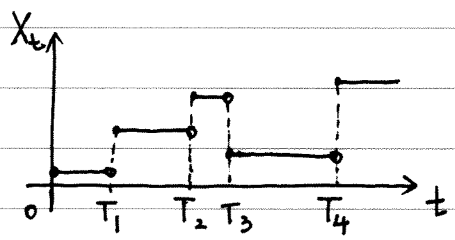
$$\begin{cases} \partial_t (u_0^+ + u_0^-) = \dfrac{\alpha^2}{2\beta}\dfrac{\partial^2}{\partial x^2}(u_0^+ + u_0^-) \\ (u_0^+ + u_0^-)\big|_{t=0} = 2f \end{cases}$$

3. Chemical reaction kinetics:

1) Definition.

Well-stirred system having $N$ molecular species $\{S_1, \cdots, S_N\}$ interacting through $M$ reaction channels $\{R_1, \cdots, R_M\}$

Jump process:




State of the system: $\vec{X}_t = (X_t^1, \cdots, X_t^N) \in \mathbb{N}^N$

$X_t^k$: # of molecules for $k$-th specie : $k=1,\cdots,N$

Reaction $R_j$ characterized by rate function $a_j(\vec{x})$ and state change vector

$$\vec{\nu}_j = (\nu_j^1, \cdots, \nu_j^N) \qquad a_j(\vec{x}) \geq 0$$

$a_j(\vec{x})dt$: the probability that the system experience $R_j$ reaction in time $dt$.

$$\vec{X}_t \longrightarrow \vec{X}_t + \vec{\nu}_j \quad \text{if } j\text{th reaction fires}$$

$$a_0(\vec{x}) \triangleq \sum_{j=1}^{M} a_j(\vec{x})$$

2) Chemical master eqn. (Forward Kolmogorov eqn)

$$p(\vec{x}, t+dt) = \sum_{j=1}^{M} p(\vec{x}-\vec{\nu}_j, t) \cdot a_j(\vec{x}-\vec{\nu}_j)\,dt$$

$$+ p(\vec{x},t)\left(1-\sum_{j=1}^{M} a_j(\vec{x})dt\right) + h.o.t.$$

$$\Rightarrow \partial_t p = \sum_{j=1}^{M} a_j(\vec{x}-\vec{\nu}_j)\, p(\vec{x}-\vec{\nu}_j, t) - \sum_{j=1}^{M} a_j(\vec{x})\, p(\vec{x},t)$$

3) Large volume scaling:

<u>Rescale process</u>: $X_t^V \in \mathbb{N}^N/V \sim O(1)$ $V$: Volume size $V \gg 1$

Jump size $\vec{\nu}_j/V$, jump rate $a_j(V\cdot\vec{x}) \sim O(V)$

Assume $a_j(V\vec{x}) = V\cdot \tilde{a}_j(\vec{x}) + h.o.t.$ $\quad \vec{x} \sim O(1)$

<u>Def</u>: $u^\varepsilon(x,t) = \mathbb{E}^x f(\vec{X}_t^V)$, $\quad \varepsilon \triangleq 1/V$

$$\partial_t u^\varepsilon = \sum_{j=1}^{M} a_j\left(\tfrac{1}{\varepsilon}\vec{x}\right)\left[u^\varepsilon(\vec{x}+\varepsilon\vec{\nu}_j) - u^\varepsilon(\vec{x})\right]$$

$$u^\varepsilon = u_0 + \varepsilon u_1 + \cdots$$

$$\partial_t u_0 = \sum_{j=1}^{M} \tilde{a}_j(\vec{x})\, \vec{\nu}_j \cdot \nabla u_0 = \bar{\mathcal{L}} u_0$$

$\bar{\mathcal{L}}$: generator corresponds to Reaction Rate Eqn.

This is a simplest version of multiscale dynamics.

<u>HW</u>: For a strongly continuous contraction semigroup $S(t)$ generated by $\mathcal{L}$ on Banach space $B$, i.e. $\|S(t)\| \leq 1$,

and $\lim_{t\to 0+} \|S(t)f - f\| = 0 \quad \forall f \in B$.

Assume that $S(t) \to Q$ as $t \to +\infty$, then

a) $Q^2 = Q$ and $Q$ is a contraction operator.

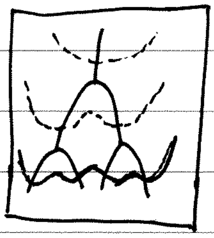
b) $S(t)Q = QS(t), \quad \forall t \geq 0$.

c) $R(Q) = N(\mathcal{L})$,

d) $N(Q) = \overline{R(\mathcal{L})}$.

# Lect 2 Energy Landscape and Quasipotential

## A. Waddington's energy landscape: (1957)



Cell differentiation: epigenetic landscape

J. Wang's landscape theory: $\frac{dx}{dt} = b(x)$ ODE.

step1: Add noise:

$$dX_t = b(X_t)dt + \sqrt{2D}\,dW_t \qquad W_t \sim \text{B.M.}$$

step2: Potential landscape:

$$\partial_t p + \nabla\cdot J = 0 \qquad J = b(x)p - D\nabla p \qquad \text{(Fokker-Planck)}$$

$t \to +\infty \downarrow$

steady state: $p_{ss}(x)$. $\nabla\cdot J_{ss} = 0$

Potential: $U(x) \triangleq -D\ln p_{ss}(x)$

step3: Decomposition:

$$b = \frac{1}{p_{ss}} D\nabla p_{ss} + \frac{J_{ss}}{p_{ss}} = \underbrace{-\nabla U}_{\text{Gradient force}} + \underbrace{\frac{J_{ss}}{p_{ss}}}_{\text{"Curl force"}}$$

Question: Connection with quasipotential in Large deviation theory?

Freidlin-Wentzell: Random perturbations of dynamical systems, Springer, 1979.

B. Large deviation theory and quasi-potential:

1. LDT for i.i.d. R.V.s :

$\{X_k\}_{k=1}^n$ i.i.d. $\bar{S}_n \triangleq \frac{1}{n}\sum_{k=1}^n X_k$ , $\mu \triangleq \mathbb{E}X_k$, $\sigma^2 = \mathbb{E}(X_k-\mu)^2$

LLN: $\bar{S}_n \to \mu \quad n \to +\infty$

CLT: $(\bar{S}_n - \mu)/\frac{\sigma}{\sqrt{n}} \Rightarrow N(0,1) \quad n \to +\infty$

LDT: $\frac{1}{n}\log \mathbb{P}(\bar{S}_n \in \Gamma) = -\inf_{x\in\Gamma} I(x)$ "regular enough $\Gamma$"

$I(x)$: rate function

Assume $X_k$ has exponential moments: $\Lambda(\lambda) \triangleq \ln \mathbb{E} e^{\lambda X_k}$.

Gärtner-Ellis:

$$I(x) = \sup_\lambda \{\lambda x - \Lambda(\lambda)\} \quad \text{Legendre transform}$$

If $X_k \sim N(0,1)$, $I(x) = x^2/2$.

2. LDT for SDEs under small perturbations.

Consider the SDEs:

$$dX_t^\varepsilon = b(X_t)dt + \sqrt{\varepsilon}\,\sigma(X_t)\,dW_t \quad \text{(Diffusion process)}$$

or

$$X_t^\varepsilon = X_0^\varepsilon + \sum_{j=1}^M \varepsilon \vec{\nu}_j P_j\left(\int_0^t a_j(\varepsilon^{-1}X_s^\varepsilon)ds\right) \quad \text{(CKS)}$$

$P_j(t)$: independent uni-rate Poisson processes, $j=1,\cdots,M$

NB: $a_j(\varepsilon^{-1}\vec{x}) = \varepsilon^{-1}\tilde{a}_j(\vec{x}) + \text{h.o.t.}$

and CKS has the chemical master eqn (for rescaled process) as its forward Kolmogorov eqn.

Under suitable conditions

a) Diffusion: $b(x)$, $\sigma(x)$ bounded, Lipschitz, $A(x)=\sigma\cdot\sigma^T(x)$ uniformly elliptic;

b) CKS: $\{\tilde{a}_j(x)\}_j$ Lipschitz, non-negative, non-escaping in $\overline{\mathbb{R}^d_+}$.

Def: $S = C[0,T]^d$ (Diffusion) or $\mathbb{D}^d[0,T]$ (CKS) : RCLL functions equipped with ↓ Uniform metric $L^\infty$ ; ↓ Skorohod metric

$X^\varepsilon$ satisfies large deviation principle:

i) Closed set $F \subset S$

$$\overline{\lim_{\varepsilon\to 0+}}\, \varepsilon \log \mathbb{P}(X^\varepsilon \in F) \le -\inf_{\varphi\in F} I_T[\varphi]$$

ii) Open set $G \subset S$

$$\underline{\lim_{\varepsilon\to 0+}}\, \varepsilon \log \mathbb{P}(X^\varepsilon \in G) \ge -\inf_{\varphi\in G} I_T[\varphi]$$

where

$$I_T[\varphi] = \begin{cases} \int_0^T L(\varphi,\dot{\varphi})\,dt, & \varphi \text{ absolutely continuous} \\ +\infty, & \text{otherwise} \end{cases}$$

is the action functional.

$L(x,y)$: Lagrangian (in classical mechanics)

Diffusion case:

$$L(x,y) = \frac{1}{2}\| y - b(x)\|_A^2 \qquad \|x\|_A^2 \triangleq x^T \cdot A^{-1} \cdot x$$

$$H(x,p) \triangleq \sup_y \{ py - L(x,y)\} \qquad \text{Hamiltonian}$$

$$\Rightarrow H(x,p) = \frac{1}{2} p^T \cdot A(x) \cdot p + b(x) \cdot p$$

CKS:

$$H(x,p) = \sum_{j=1}^{M} a_j(x) \left( e^{\vec{p} \cdot \vec{\nu}_j} - 1 \right)$$

L does NOT have closed form!

C. Quasi-potential (QP):

1. Take $x$ as a stationary point of $\dot{x} = b(x)$, i.e. $b(x) = 0$

and $Re(\lambda(J(x))) < 0$, $J(x) = \nabla b(x)$ : Jacobian.

Def

$$W(y,t;x) \triangleq \inf_{\substack{\varphi(0)=x \\ \varphi(t)=y}} I_t[\varphi]$$

and

$$V(y;x) \triangleq \inf_{t \geq 0} W(y,t;x) = \inf_{t\geq 0} \inf_{\substack{\varphi(0)=x \\ \varphi(t)=y}} I_t[\varphi]$$

$\uparrow$

Local QP starting from $x$.

2. Variational calculus:

Take $\varphi(0) = x$, $\varphi(t) = y$; $x$ fixed, $y$ may be varied.

$\varphi(\cdot)$: minimizing path of $I_t[\varphi]$

$$\delta W = \frac{\partial L}{\partial \dot\varphi}\delta\varphi\Big|_0^t + \int_0^t \underbrace{\left(\frac{\partial L}{\partial \varphi} - \frac{d}{ds}\left(\frac{\partial L}{\partial \dot\varphi}\right)\right)}_{=0 \text{ for minimizing path}} \delta\varphi\, ds$$

$\delta\varphi(0)=0$, $p \triangleq \frac{\partial L}{\partial \dot\varphi}$ (Legendre transform)

$$\Rightarrow \delta W = p\,\delta y \Rightarrow p = \frac{\partial W}{\partial y}, \text{ i.e. } p = \nabla_y W$$

Along the path,

$$\left.\begin{aligned} \frac{dW}{dt} &= L(\varphi,\dot\varphi) \\ \frac{dW}{dt} &= \partial_t W + \frac{\partial W}{\partial y}\dot\varphi \end{aligned}\right\} \Rightarrow \partial_t W + (p\dot\varphi - L) = 0$$

$$\Rightarrow \partial_t W + H(\varphi, \nabla_y W) = 0 \qquad \text{HJE}$$

For $V(y;x)$, $\quad H(y, \nabla_y V) = 0$

3. Global vs Local QP:

1) Why "quasi"-potential? $\quad W(y)$: QP

$$dX_t = b(X_t)dt + \sqrt{\varepsilon}\,\sigma(X_t)dW_t$$

Take $b(x) = -\nabla U(x)$, $\sigma = 1$

$$H(x, \nabla W) = 0 \Rightarrow \frac{1}{2}|\nabla W|^2 - \nabla U \cdot \nabla W = 0$$

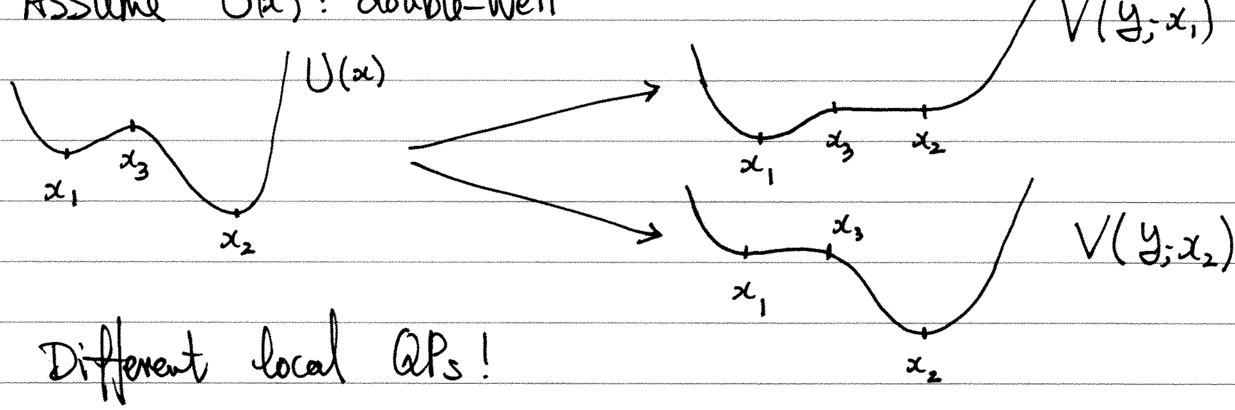
$$\Rightarrow W = 2U$$

$W(y)$ plays the role of potential.

2) Local QPs: $\quad b(x) = -\nabla U(x), \quad \sigma(x) = 1$

$$I_t[\varphi] = \frac{1}{2}\int_0^t |\dot{\varphi} + \nabla U(\varphi)|^2 ds$$

Assume $U(x)$: double-well




Different local QPs!

3) Global QP: $\quad W(y)$

$$W(y) \triangleq \lim_{\varepsilon \to 0+} -\varepsilon \log P_{ss}^{\varepsilon}(y|x) = \lim_{\varepsilon \to 0+} \lim_{t \to \infty} -\varepsilon \log P^{\varepsilon}(y,t|x,0)$$

$$\neq \lim_{t \to \infty} \lim_{\varepsilon \to 0+} -\varepsilon \log P^{\varepsilon}(y,t|x,0)$$

$$= \lim_{t \to \infty} W(y,t;x)$$

$$= V(y;x)$$

However, $W(y)$ can be obtained from local QPs by pruning and sticking procedure.

See Zhou-Li, J. Chem. Phys. 144 (2016), 094109 for more details on different proposals about energy landscapes.

4) Facts on global and local QP:

2.7

- Global QP characterize the long time equilibrium behavior
  Local QP 〃 the difficulty of local transitions.

- Transitivity does not hold in general for non-gradient system:

$$P_{a\to b} > P_{b\to a},\quad P_{b\to c} > P_{c\to b} \not\Rightarrow P_{a\to c} > P_{c\to a}$$

5) Lyapunov property of $W$:

Def $W = 2U$, then if $U \in C^1(\Omega)$, $\Omega \subset \mathbb{R}^d$

$$\tilde{H}(x, \nabla U) = 0,\qquad \tilde{H}(x,p) = p^T A p + b\cdot p$$

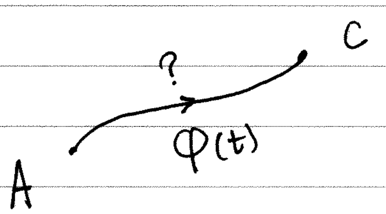
Def: $l(x) = b(x) + A(x)\cdot \nabla U(x)$

$$\Rightarrow\ l(x)\cdot \nabla U = 0$$

For ODE $\dot{x} = b(x)$.

$$\frac{dU}{dt} = \nabla U\cdot b = \nabla U(-A\nabla U + l) = -\nabla U\cdot A\cdot \nabla U \leqslant 0$$

6) Analysis of uphill and downhill path:

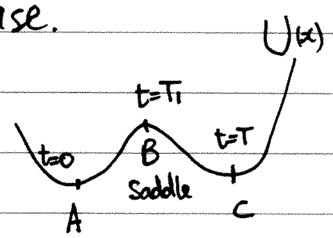


A, C are two stable fixed points.

Transition pathway from A to C?

$$I_T[\varphi] = \int_0^{T_1} \|\dot{\varphi} - A\cdot\nabla U \pm l\|_A^2\, dt + 4\underbrace{\int_0^{T_1} \dot{\varphi}\cdot \nabla U\, dt}_{=\,4\Delta U} + \int_{T_1}^{T} \|\dot{\varphi} + A\cdot\nabla U - l\|_A^2\, dt \geqslant 4\Delta U \qquad (W = 2U)$$

U: play the role of potential in the gradient case.




minimizing path:

$$\frac{d\varphi}{dt} = \underbrace{A(\varphi)\cdot\nabla U(\varphi)}_{\text{steepest ascent part}} + l(\varphi), \quad 0 \le t \le T_1$$

$$\frac{d\varphi}{dt} = -A(\varphi)\cdot\nabla U(\varphi) + l(\varphi), \quad T_1 \le t \le T$$

$$= b(\varphi) \quad \text{steepest descent part}$$

Effect of $l$: non-gradient force!

<u>NB</u>: a) $T_1, T$ should be $+\infty$ in final minimization.

b) Minimizing path is not smooth at the saddle point for non-gradient system due to the presence of $l(x)$.

<u>Open issue</u>: How to construct the desired low-dim landscape for a high-dim dynamics?

<u>HW1</u>: Explore the minimizing path for the CKS dynamics.

<u>HW2</u>: Explore the relation between the potential landscape from $P_{ss}(x)$ and QP $W(x)$.

<u>Ref</u>: Lv. et al, PLoS One 9 (2014), e88167.

Lv et al. PLoS CB 11 (2015), e1004156.

# Lect 3 Two-scale LDT for CKS

## A. Motivating example:

### 1. Genetic switching with positive feedback:

$$
\begin{array}{ccccc}
\mathrm{DNA}_{in} & & \phi & & \phi \\
G(Z_2)\uparrow \;\downarrow F(Z_2) & & \uparrow \gamma & & \uparrow 1 \\
\mathrm{DNA}_{act} & \xrightarrow{a} & \mathrm{mRNA} & \xrightarrow{\gamma b} & \mathrm{Protein} \\
 & & Z_1 & & Z_2
\end{array}
$$

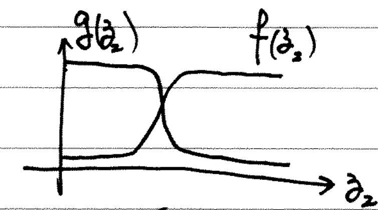



Fig. 1

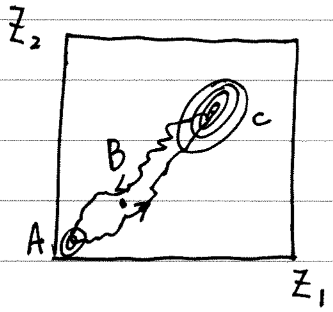
$(Z_1, Z_2)$ : # of molecules for mRNA and Protein $\sim O(\varepsilon^{-1})$

$\vec{z} \triangleq (z_1, z_2) = \varepsilon (Z_1, Z_2)$ : rescaled variable.

$F(Z_2) = \varepsilon^{-1} f(z_2)$, $G(Z_2) = \varepsilon^{-1} g(z_2)$ positive feedback (in Fig. 1)

$a \sim \varepsilon^{-1} b^{-1}$, $b \sim O(1)$, $\gamma \sim O(1)$

$D \in \{0, 1\}$ : DNA state. $\begin{cases} D=0: & \text{Inactive state} \\ D=1: & \text{Active state} \end{cases}$




With suitable choice of $f, g$, $\exists$ two metastable regims around A and C:

Rare transitions between A and C can be characterized by LDT.

<u>NB</u>: DNA state D can NOT be rescaled!

### 2. Large volume limit (LLN, $\varepsilon \to 0$):

For rescaled process $(D(t), \vec{z}(t))$, define

$$h(D, \vec{z}, t) = \mathbb{E}^{(D, \vec{z})} f(D(t), \vec{z}(t))$$

and denote

$$\begin{pmatrix} u(\vec{z},t) \\ v(\vec{z},t) \end{pmatrix} = \begin{pmatrix} h(0,\vec{z},t) \\ h(1,\vec{z},t) \end{pmatrix}.$$

Then

$$\frac{d}{dt}\begin{pmatrix} u(\vec{z},t) \\ v(\vec{z},t) \end{pmatrix} = \mathcal{L}^{\varepsilon}\begin{pmatrix} u(\vec{z},t) \\ v(\vec{z},t) \end{pmatrix} \qquad ①$$

$\mathcal{L}^{\varepsilon}$: infinitesimal generator of the rescaled process.

$$\mathcal{L}^{\varepsilon}\begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{\varepsilon}\begin{pmatrix} -f(z_2) & f(z_2) \\ g(z_2) & -g(z_2) \end{pmatrix}\cdot\begin{pmatrix} u \\ v \end{pmatrix} + A^{\varepsilon}\begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{1}{\varepsilon} b^{-1}\delta^{\varepsilon}_{z_1} v \end{pmatrix} \qquad ②$$

where $\delta^{\varepsilon}_{z_1} v \triangleq v(z_1+\varepsilon, z_2) - v(z_1, z_2)$ : transcription process,

and $A^{\varepsilon}$ characterizes the translation and decay processes:

$$A^{\varepsilon} g \triangleq \Big(\underbrace{\varepsilon^{-1} z_2 \delta^{-\varepsilon}_{z_2} + \gamma\varepsilon^{-1} z_1 \delta^{-\varepsilon}_{z_1}}_{\text{decay}} + \underbrace{\gamma b \varepsilon^{-1} z_1 \delta^{\varepsilon}_{z_2}}_{\text{translation}}\Big) g$$

② has the form:

$$\mathcal{L}^{\varepsilon} = \frac{1}{\varepsilon}\mathcal{L}_1 + \mathcal{L}_2 + \text{h.o.t.}$$

$$\mathcal{L}_1 = \begin{pmatrix} -f(z_2) & f(z_2) \\ g(z_2) & -g(z_2) \end{pmatrix}.$$

$$\mathcal{L}_2 \begin{pmatrix} u \\ v \end{pmatrix} = \mathcal{A} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ b^{-1}\partial_{z_1} v \end{pmatrix}$$

$$\mathcal{A} g \triangleq \left(-z_2 \partial_{z_2} - \gamma z_1 \partial_{z_1} + \gamma b z_1 \partial_{z_2}\right) g$$

Let $\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} + \varepsilon \begin{pmatrix} u_1 \\ v_1 \end{pmatrix} + \text{h.o.t.}$

$O(\varepsilon^{-1})$: $\quad u_0(\vec{z}) = v_0(\vec{z})$

$O(1)$: $\quad \dfrac{du_0}{dt} = (\gamma b z_1 - z_2)\partial_{z_2} u_0 + \left(b^{-1}\dfrac{f}{f+g} - \gamma z_1\right)\partial_{z_1} u_0$

This corresponds to the deterministic process

$$\begin{cases} \dfrac{dz_1}{dt} = \dfrac{b^{-1} f(z_2)}{f(z_2)+g(z_2)} - \gamma z_1 \\ \dfrac{dz_2}{dt} = \gamma b z_1 - z_2 \end{cases} \quad : \text{mean field ODEs.}$$

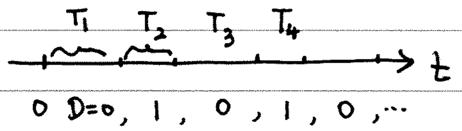
B. Two-scale large deviation:

For the motivating example, DNA state has fast switching instead of simple large volume scaling. We need LDT for multiscale dynamics. Pioneering works include Veretennikov, Liptser, et al.

1. LDT for fast process:

The key of the LDT for fast process is to treat the so-called <u>occupation time</u> instead of states.

1)




Suppose the slow variable $z_2$ is fixed, for fast process $D(t)$:

Occupation measure:

$$M_t(i) \triangleq \frac{1}{t}\int_0^t \delta_i(D(s))ds, \qquad i \in \{0,1\}$$

$$\delta_i(j) \triangleq \delta_{ij}, \quad i,j = 0,1. \qquad \text{Kronecker's } \delta\text{-function.}$$

$M_t(i)$: the fraction of time that $D(t)$ spends at $i$.

Obviously, LLN: $M_t(i) \xrightarrow{t\to\infty} \pi_i$ invariant distribution.

2) To understand the LDT for $M_t$, consider the discrete time case for simplicity.

K-state Markov chain with transition matrix $P = (P_{ij})_{i,j=1}^K$

Occupation measure:

$$L_{n,i} \triangleq \frac{1}{n}\sum_{k=1}^{n} \delta_i(X_k) \qquad X_k \in \{1,2,\cdots,K\}$$

Gärtner–Ellis: For ergodic $P$,

$$\Lambda(s) \triangleq \lim_{n\to\infty} \frac{1}{n}\ln\langle e^{ns\cdot L_n}\rangle \qquad s \in \mathbb{R}^K$$

then

$$I(w) = \sup_s \{s\cdot w - \Lambda(s)\} \qquad w \in \mathcal{M}_1: \text{prob. measure}$$

Here

$$\langle e^{nS\cdot L_n}\rangle = \sum_{i_1\cdots i_n} \mu_{i_1} e^{S_{i_1}} p_{i_1 i_2} e^{S_{i_2}} \cdots p_{i_{n-1} i_n} e^{S_{i_n}}$$

$\mu$: initial distribution

Def:

$$(P_s)_{ij} \triangleq p_{ij} e^{S_j}, \quad (\mu_s)_i \triangleq \mu_i e^{S_i}$$

$$\langle e^{nS\cdot L_n}\rangle = \mu_s^T \cdot (P_s)^{n-1} \cdot 1 \qquad 1: \text{vector with each component}=1$$

If $P$ is ergodic (irreducible), $P_s$ is also ergodic

$$\rho_s \triangleq \max |\lambda(P_s)|, \quad \text{spectral radius}: \in \mathbb{R}^+, \text{ simple}$$

$$\Rightarrow \Lambda(s) = \rho_s \qquad\qquad w \in \mathcal{M}_1$$

and

$$I(w) = \sup_s \{ S\cdot w - \ln \rho_s \} \qquad (3)$$

Furthermore, it can be shown that (c.f. Dembo-Zeitouni; LDT and Applicatio

$$I(w) = \sup_{u>0} \Big\{ \underbrace{-\sum_{i=1}^{K} w_i \ln \frac{(uP)_i}{u_i}}_{\triangleq\, H(w,u)\ :\ \text{Hamiltonian}} \Big\} \qquad (4)$$

3) The above result holds similarly for ergodic Q-process:

$$I(w) = \sup_{u>0} \Big\{ -\sum_i w_i \frac{(Qu)_i}{u_i} \Big\}$$

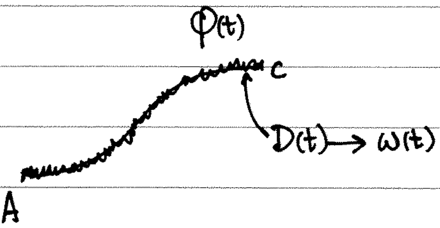
$Q$: generator

4) Two-scale LDT:

The fast and slow scales are fully coupled:

LDT: $\mathbb{P}(\vec{z}_\varepsilon, \nu_\varepsilon \in \Gamma) \asymp e^{-\varepsilon^{-1} \inf_{(\varphi,\omega)\in\Gamma} I(\varphi,\omega)}$

$$I(\varphi,\omega) = I_s(\varphi,\omega) + I_f(\varphi,\omega)$$




For fast variables, given $\varphi$

$$I_f(\varphi,\omega) = \int_0^T L_f(\varphi,\omega)\,dt$$

Depends on $\varphi$.

$$L_f(\varphi,\omega) = \sup_{u>0}\left\{ -\sum_i \omega_i \frac{(Q(\varphi)u)_i}{u_i} \right\}$$

For slow variables, given $\omega$

$$I_s(\varphi,\omega) = \int_0^T L_s(\varphi,\dot{\varphi},\omega)\,dt$$

$$L_s(\varphi,\dot{\varphi},\omega) = \sup_p \{ \dot{\varphi}\cdot p - H_s(\varphi,p,\omega) \}$$

$$H_s(\varphi,p,\omega) = \sum_{k=1}^{K} \omega_k \sum_{j=1}^{M} a_j(\varphi)\left( e^{p\cdot\nu_j} - 1 \right)$$

Reactions depend on state $k$.

The above formulation holds because the switching and translation in the same timescale (primitive variable)

5) Extension:

Different time scalings:

$$\begin{tikzcd}
\text{DNA}_{in} \arrow[d, shift right, "f(n)"'] \arrow[dr, "k_1"] & & \\
\text{DNA}_{act} \arrow[u, shift right, "g(n)"'] \arrow[r, "k_1"'] & \underset{n}{\text{Protein}} \arrow[r, "\gamma"] & \phi
\end{tikzcd}$$

$k_0, k_1 \gg \gamma \sim O(1)$

$f, g \gg 1$

Regimes: $\begin{cases} k_0, k_1 \gg f, g \\ k_0, k_1 \sim f, g \\ k_0, k_1 \ll f, g \end{cases}$ $\rightarrow$ Different LDT results!

<u>Ref.</u> Lin-Li, CMS 15 (2017), 123.

Lin-Li, J. Phys. A 49 (2016), 135204.

<u>HW</u>: Try to prove ③ $\Leftrightarrow$ ④.

( Utilize a result of Perron-Frobenius Thm:

$P$ is irreducible, non-negative, then for $\forall i \in \{1, 2, \cdots, K\}$

$\forall u$ s.t. $u_j > 0$ for $\forall j$,

$$\lim_{n\to\infty} \frac{1}{n} \log \sum_j (P^n)_{ij} u_j = \lim_{n\to\infty} \frac{1}{n} \log \sum_j u_j (P^n)_{ji} = \log \rho(P) \text{ .)}$$